GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: June 30, 2001, 04:16:36 ; Search time 2344.34 Seconds
(without alignments)
11407.790 Million cell updates/sec

Title: US-09-403-262A-1
Perfect score: 1729
Sequence: 1 cctctgtccttccccccagg.....cttatttttatttttaatta 1729

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_ba3:*
4: gb_in1:*
5: gb_in2:*
6: gb_in3:*
7: gb_om:*
8: gb_ov:*
9: gb_pat1:*
10: gb_pat2:*
11: gb_ph:*
12: gb_pl1:*
13: gb_pl2:*
14: gb_pl3:*
15: gb_pl4:*
16: em_ba1:*
17: em_ba2:*
18: em_fun:*
19: em_htgo_hum:*
20: em_htgo_inv:*
21: em_htgo_rod:*
22: em_htg_hum1:*
23: em_htg_hum2:*
24: em_htg_hum3:*
25: em_htg_hum4:*
26: em_htg_hum5:*
27: em_htg_hum6:*
28: em_htg_hum7:*
29: em_htg_hum8:*
30: em_htg_inv1:*
31: em_htg_inv2:*
32: em_htg_other:*
33: em_htg_rod:*
34: em_hum1:*
35: em_hum2:*
36: em_hum3:*
37: em_hum4:*
38: em_hum5:*
39: em_hum6:*
40: em_hum7:*
41: em_in:*
42: em_om:*
43: em_or:*
44: em_ov:*
45: em_pat:*
46: em_ph:*
47: em_pl:*
48: em_ro:*
49: em_sts:*
50: em_sy:*
51: em_un:*
52: em_vi:*
53: gb_sts1:*
54: gb_sts2:*
55: gb_sts3:*
56: gb_sy:*
57: gb_un:*
58: gb_vi1:*
59: gb_vi2:*
60: gb_htg1:*
61: gb_htg2:*
62: gb_htg3:*
63: gb_htg4:*
64: gb_htg5:*
65: gb_htg6:*
66: gb_htg7:*
67: gb_htg8:*
68: gb_htg9:*
69: gb_htg10:*
70: gb_htg11:*
71: gb_htg12:*
72: gb_htg13:*
73: gb_htg14:*
74: gb_htg15:*
75: gb_htg16:*
76: gb_htg17:*
77: gb_htg18:*
78: gb_htg19:*
79: gb_htg20:*
80: gb_htg21:*
81: gb_htg22:*
82: gb_htg23:*
83: gb_htg24:*
84: gb_htg25:*
85: gb_pr1:*
86: gb_pr2:*
87: gb_pr3:*
88: gb_pr4:*
89: gb_pr5:*
90: gb_pr6:*
91: gb_pr7:*
92: gb_pr8:*
93: gb_pr9:*
94: gb_ro1:*
95: gb_ro2:*
96: gb_in4:*
97: gb_pr10:*
98: em_ba3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1729	100.0	1729	9	A84072	A84072 Sequence 1
	2	1729	100.0	1729	12	AF098674	AF098674 Lycopersi
c	3	1729	100.0	57047	14	LES303345	AJ303345 Lycopersi
	4	1223.2	70.7	1296	9	A84080	A84080 Sequence 9
c	5	294.2	17.0	87967	12	AC005223	AC005223 Arabidops
	6	294.2	17.0	109171	12	AC002328	AC002328 Genomic s
	7	275.2	15.9	27056	14	CRU303349	AJ303349 Capsella
	8	218.8	12.7	687	9	A84084	A84084 Sequence 13

	9	141.6	8.2	2210	13	ATRGA1 Y11336 A.thaliana
	10	140	8.1	2201	14	ATY15194 Y15194 Arabidopsis
c	11	140	8.1	95137	12	AC005560 AC005560 Arabidops
	12	133.6	7.7	1951	13	ATRGA2 Y11337 A.thaliana
	13	133.6	7.7	1964	9	A64697 A64697 Sequence 1
	14	133.6	7.7	1964	14	ATY15193 Y15193 Arabidopsis
	15	133.6	7.7	132699	12	AC006917 AC006917 Genomic s
	16	132	7.6	82289	13	ATAC009895 AC009895 Arabidops
	17	125.4	7.3	1779	10	AX081276 AX081276 Sequence
	18	123.8	7.2	1779	10	AX081278 AX081278 Sequence
	19	115.4	6.7	1642	9	A64705 A64705 Sequence 9
	20	115.4	6.7	1643	9	A64699 A64699 Sequence 3
	21	105.8	6.1	1359	12	AF036300 AF036300 Arabidops
	22	105.8	6.1	2130	13	AF210731 AF210731 Arabidops
	23	105.8	6.1	4081	13	ATA224957 AJ224957 Arabidops
	24	105.8	6.1	85702	12	AC020665 AC020665 Arabidops
	25	105.8	6.1	108365	12	AC015447 AC015447 Arabidops
	26	105.4	6.1	1642	9	A64701 A64701 Sequence 5
	27	104.8	6.1	69850	12	AB010700 AB010700 Arabidops
c	28	104.4	6.0	26604	13	ATK3M16 AL391150 Arabidops
c	29	102.2	5.9	100512	12	AC010795 AC010795 Arabidops
c	30	100.6	5.8	92745	12	AC079279 AC079279 Arabidops
	31	100.6	5.8	105733	12	AC012561 AC012561 Arabidops
c	32	98.4	5.7	84196	14	ATT3A5 AL132979 Arabidops
	33	92.6	5.4	1636	9	A64703 A64703 Sequence 7
	34	92.2	5.3	1065	12	AF036302 AF036302 Arabidops
	35	86.2	5.0	1793	13	AF153443 AF153443 Arabidops
c	36	86.2	5.0	59372	12	AB023039 AB023039 Arabidops
	37	84.8	4.9	1078	12	AF036308 AF036308 Arabidops
	38	84.6	4.9	1231	12	AF036301 AF036301 Arabidops
	39	82.4	4.8	104830	12	AC022464 AC022464 Genomic s
	40	81.8	4.7	83689	12	AB017067 AB017067 Arabidops
	41	80	4.6	1634	13	AF210732 AF210732 Arabidops
c	42	80	4.6	96654	12	AC007211 AC007211 Arabidops
	43	77.4	4.5	3663	12	AB012639 AB012639 Nicotiana
c	44	76	4.4	198788	13	ATCHRIV46 AL161546 Arabidops
	45	74.4	4.3	7218	10	I66494 I66494 Sequence 14

ALIGNMENTS

```
RESULT   1
A84072
LOCUS       A84072       1729 bp    DNA              PAT       21-JAN-2000
DEFINITION  Sequence 1 from Patent WO9846759.
ACCESSION   A84072
VERSION     A84072.1  GI:6733213
KEYWORDS    .
SOURCE      tomato.
  ORGANISM  Lycopersicon esculentum
            Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
            Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids
            I; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE   1  (bases 1 to 1729)
  AUTHORS   Theres,N.
  TITLE     PLANTS WITH CONTROLLED SIDE-SHOOT FORMATION AND/OR CONTROLLED
            ABSCISSION AREA FORMATION
  JOURNAL   Patent: WO 9846759-A 1 22-OCT-1998;
            THERES NIKOLAUS (DE)
FEATURES             Location/Qualifiers
     source          1. .1729
                     /organism="Lycopersicon esculentum"
                     /db_xref="taxon:4081"
BASE COUNT      487 a    372 c    294 g    576 t
ORIGIN


  Query Match             100.0%;  Score 1729;  DB 9;  Length 1729;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1729;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 cctctgtccttccccccaggtccccttttttttccttttctctctctccttttatttctcttt 60
```

```
Db      1 CCTCTGTCCTTCCCCCCAGGTCCCCTTTTTTTCCTTTCTCTCTCTCCTTTATTTCTCTTT 60

Qy     61 tcataagcatattctttctctctctagggttttcactttcacctgaaatagtgttgttaa 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TCATAAGCATATTCTTTCTCTCTCTAGGGTTTTCACTTTCACCTGAAATAGTGTTGTTAA 120

Qy    121 attgaatgatatgttaggatcctttggttcttcatcatctcaatctcaccctcatcatga 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 ATTGAATGATATGTTAGGATCCTTTGGTTCTTCATCATCTCAATCTCACCCTCATCATGA 180

Qy    181 tgaagaatcttctgatcatcatcaacagcgtagattcaccgctactgctacaactatcac 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TGAAGAATCTTCTGATCATCATCAACAGCGTAGATTCACCGCTACTGCTACAACTATCAC 240

Qy    241 caccaccaccatcactacctcaccagctattcaaatccgccagctactcattagctgtgc 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CACCACCACCATCACTACCTCACCAGCTATTCAAATCCGCCAGCTACTCATTAGCTGTGC 300

Qy    301 ggagttgatttcgcagtccgatttctcggccgcgaaaagactccttactatattatcaac 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GGAGTTGATTTCGCAGTCCGATTTCTCGGCCGCGAAAAGACTCCTTACTATATTATCAAC 360

Qy    361 taactcatctccttttggtgattcaactgaacggttagtccatcaatttactcgcgcact 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 TAACTCATCTCCTTTTGGTGATTCAACTGAACGGTTAGTCCATCAATTTACTCGCGCACT 420

Qy    421 ttcccttcgtctcaaccgctatatatcgtcaaccaccaatcatttcatgacacctgttga 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 TTCCCTTCGTCTCAACCGCTATATATCGTCAACCACCAATCATTTCATGACACCTGTTGA 480

Qy    481 aacaactccaactgattcttcttcttcgtcatcattagctctaattcaatcatcatatct 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 AACAACTCCAACTGATTCTTCTTCTTCGTCATCATTAGCTCTAATTCAATCATCATATCT 540

Qy    541 atctctaaaccaagttacccctttcataaggtttactcaattaaccgctaatcaagcgat 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 ATCTCTAAACCAAGTTACCCCTTTCATAAGGTTTACTCAATTAACCGCTAATCAAGCGAT 600

Qy    601 tttagaagcgattaacggtaatcatcaagcaatccacatcgttgatttcgacattaatca 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 TTTAGAAGCGATTAACGGTAATCATCAAGCAATCCACATCGTTGATTTCGACATTAATCA 660

Qy    661 cggggttcaatggccaccgttaatgcaagcactagctgatcgttaccctgctcccactct 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 CGGGGTTCAATGGCCACCGTTAATGCAAGCACTAGCTGATCGTTACCCTGCTCCCACTCT 720

Qy    721 tcgaatcaccggtactggaaatgaccttgataccttcgtagaacaggtgatcgtttagc 780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 TCGAATCACCGGTACTGGAAATGACCTTGATACCCTTCGTAGAACAGGTGATCGTTTAGC 780

Qy    781 taaatttgctcactcattagggttgagatttcaattccatcctctttatatagccaataa 840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 TAAATTTGCTCACTCATTAGGGTTGAGATTTCAATTCCATCCTCTTTATATAGCCAATAA 840

Qy    841 taaccacgatcacgatgaagatccttctattatttcctccattgtactactccctgatga 900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 TAACCACGATCACGATGAAGATCCTTCTATTATTTCCTCCATTGTACTACTCCCTGATGA 900

Qy    901 aaccctagctatcaactgtgttttctacctccaccgccttttaaaagaccgcgaaaagtt 960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 AACCCTAGCTATCAACTGTGTTTTCTACCTCCACCGCCTTTTAAAAGACCGCGAAAAGTT 960

Qy    961 aaggatttttttgcatagggttaagtcaatgaaccctaaaattgttacaatcgcggagaa 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 AAGGATTTTTTTGCATAGGGTTAAGTCAATGAACCCTAAAATTGTTACAATCGCGGAGAA 1020

Qy   1021 ggaagcaaatcataaccatcctctttttttacaaagattcatcgaggcgttggattatta 1080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 GGAAGCAAATCATAACCATCCTCTTTTTTTACAAAGATTCATCGAGGCGTTGGATTATTA 1080

Qy   1081 tacagctgtgtttgattcactggaagctacattgccaccgggtagtcgagagaggatgac 1140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    1081 TACAGCTGTGTTTGATTCACTGGAAGCTACATTGCCACCGGGTAGTCGAGAGAGGATGAC 1140

Qy    1141 agttgaacaagtgtggtttgggagagagattgttgatatcgttgcgatggaaggagataa 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 AGTTGAACAAGTGTGGTTTGGGAGAGAGATTGTTGATATCGTTGCGATGGAAGGAGATAA 1200

Qy    1201 aaggaaagaaagacatgaaaggtttagatcatgggaagttatgttgaggagttgtggatt 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 AAGGAAAGAAAGACATGAAAGGTTTAGATCATGGGAAGTTATGTTGAGGAGTTGTGGATT 1260

Qy    1261 tagtaatgttgctttaagccctttgcattatcacaagctaagcttctttgagacttca 1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 TAGTAATGTTGCTTTAAGCCCTTTTGCATTATCACAAGCTAAGCTTCTTTTGAGACTTCA 1320

Qy    1321 ttatccttctgaaggctatcaactcggagtttcgagtaattctttcttcttaggttggca 1380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321 TTATCCTTCTGAAGGCTATCAACTCGGAGTTTCGAGTAATTCTTTCTTCTTAGGTTGGCA 1380

Qy    1381 aaatcaaccccttttctccatctcgtcttggcgttgagaaaaactatcaaatagccaact 1440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1381 AAATCAACCCCTTTTCTCCATCTCGTCTTGGCGTTGAGAAAAACTATCAAATAGCCAACT 1440

Qy    1441 tcagagggtaattaagactactgatagtttaggagggatctgaagaaaacgcgtggagtg 1500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1441 TCAGAGGGTAATTAAGACTACTGATAGTTTAGGAGGGATCTGAAGAAAACGCGTGGAGTG 1500

Qy    1501 aaaaccctaaataaccagattttctaatgaagttgtagtagtagaaatttgcatggtgaa 1560
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1501 AAAACCCTAAATAACCAGATTTTCTAATGAAGTTGTAGTAGTAGAAATTTGCATGGTGAA 1560

Qy    1561 gaacaatattgaagaggtattgaaatttcatgttttttttgttttacttattgatatgaa 1620
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1561 GAACAATATTGAAGAGGTATTGAAATTTCATGTTTTTTTTGTTTTACTTATTGATATGAA 1620

Qy    1621 tgttttaaaatttttaacatagaggactaggttgatgatatatagtatttaagttaacta 1680
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1621 TGTTTTAAAATTTTTAACATAGAGGACTAGGTTGATGATATATAGTATTTAAGTTAACTA 1680

Qy    1681 gtctttgtataacgcaagatcttgatcaacttatttttattttttaatta 1729
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db    1681 GTCTTTGTATAACGCAAGATCTTGATCAACTTATTTTTATTTTTAATTA 1729


RESULT   2
AF098674
LOCUS       AF098674     1729 bp    mRNA            PLN       17-JAN-1999
DEFINITION  Lycopersicon esculentum lateral suppressor protein (Ls) mRNA,
            complete cds.
ACCESSION   AF098674
VERSION     AF098674.1  GI:4160440
KEYWORDS    .
SOURCE      tomato.
  ORGANISM  Lycopersicon esculentum
            Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
            Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids
            I; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE   1  (bases 1 to 1729)
  AUTHORS   Schumacher,K., Schmitt,T., Rossberg,M., Schmitz,G. and Theres,K.
  TITLE     The Lateral suppressor (Ls) gene of tomato encodes a new member of
            the VHIID protein family
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (1), 290-295 (1999)
  MEDLINE   99093529
REFERENCE   2  (bases 1 to 1729)
  AUTHORS   Schumacher,K., Schmitt,T.T., Rossberg,M., Schmitz,G. and Theres,K.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-OCT-1998) Institut fuer Genetik, Universitaet zu
            Koeln, Carl-von-Linne-Weg 10, Koeln, NRW D-50829, Germany
FEATURES             Location/Qualifiers
     source          1. .1729
                     /organism="Lycopersicon esculentum"
                     /cultivar="VFNT Cherry"
                     /db_xref="taxon:4081"
                     /chromosome="7"
                     /map="between CD61 and CD65"
     gene            1. .1729
                     /gene="Ls"
     CDS             131. .1417
                     /gene="Ls"
                     /codon_start=1
                     /product="lateral suppressor protein"
                     /protein_id="AAD05242.1"
                     /db_xref="GI:4160441"
                     /translation="MLGSFGSSSSQSHPHHDEESSDHHQQRRFTATATTITTTTITTS
                     PAIQIRQLLISCAELISQSDFSAAKRLLTILSTNSSPFGDSTERLVHQFTRALSLRLN
                     RYISSTTNHFMTPVETTPTDSSSSSSLALIQSSYLSLNQVTPFIRFTQLTANQAILEA
                     INGNHQAIHIVDFDINHGVQWPPLMQALADRYPAPTLRITGTGNDLDTLRRTGDRLAK
                     FAHSLGLRFQFHPLYIANNNHDHDEDPSIISSIVLLPDETLAINCVFYLHRLLKDREK
                     LRIFLHRVKSMNPKIVTIAEKEANHNHPLFLQRFIEALDYYTAVFDSLEATLPPGSRE
                     RMTVEQVWFGREIVDIVAMEGDKRKERHERFRSWEVMLRSCGFSNVALSPFALSQAKL
                     LLRLHYPSEGYQLGVSSNSFFLGWQNQPLFSISSWR"
BASE COUNT      487 a    372 c    294 g    576 t
ORIGIN


  Query Match           100.0%;  Score 1729;  DB 12;  Length 1729;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 1729;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 cctctgtccttccccccaggtccccttttttttccttttctctctctccttttatttctcttt 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 CCTCTGTCCTTCCCCCCAGGTCCCCTTTTTTTCCTTTCTCTCTCTCCTTTATTTCTCTTT 60

Qy      61 tcataagcatattctttctctctctagggttttcactttcacctgaaatagtgttgttaa 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 TCATAAGCATATTCTTTCTCTCTCTAGGGTTTTCACTTTCACCTGAAATAGTGTTGTTAA 120

Qy     121 attgaatgatatgttaggatcctttggttcttcatcatctcaatctcaccctcatcatga 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 ATTGAATGATATGTTAGGATCCTTTGGTTCTTCATCATCTCAATCTCACCCTCATCATGA 180

Qy     181 tgaagaatcttctgatcatcatcaacagcgtagattcaccgctactgctacaactatcac 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 TGAAGAATCTTCTGATCATCATCAACAGCGTAGATTCACCGCTACTGCTACAACTATCAC 240

Qy     241 caccaccaccatcactacctcaccagctattcaaatccgccagctactcattagctgtgc 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CACCACCACCATCACTACCTCACCAGCTATTCAAATCCGCCAGCTACTCATTAGCTGTGC 300

Qy     301 ggagttgatttcgcagtccgatttctcggccgcgaaaagactccttactatattatcaac 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 GGAGTTGATTTCGCAGTCCGATTTCTCGGCCGCGAAAAGACTCCTTACTATATTATCAAC 360

Qy     361 taactcatctcctttggtgattcaactgaacggttagtccatcaatttactcgcgcact 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 TAACTCATCTCCTTTTGGTGATTCAACTGAACGGTTAGTCCATCAATTTACTCGCGCACT 420

Qy     421 ttcccttcgtctcaaccgctatatatcgtcaaccaccaatcatttcatgacacctgttga 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 TTCCCTTCGTCTCAACCGCTATATATCGTCAACCACCAATCATTTCATGACACCTGTTGA 480

Qy     481 aacaactccaactgattcttcttcttcgtcatcattagctctaattcaatcatcatatct 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 AACAACTCCAACTGATTCTTCTTCTTCGTCATCATTAGCTCTAATTCAATCATCATATCT 540

Qy     541 atctctaaaccaagttacccctttcataaggtttactcaattaaccgctaatcaagcgat 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 ATCTCTAAACCAAGTTACCCCTTTCATAAGGTTTACTCAATTAACCGCTAATCAAGCGAT 600

Qy     601 tttagaagcgattaacggtaatcatcaagcaatccacatcgttgatttcgacattaatca 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 TTTAGAAGCGATTAACGGTAATCATCAAGCAATCCACATCGTTGATTTCGACATTAATCA 660

Qy     661 cggggttcaatggccaccgttaatgcaagcactagctgatcgttaccctgctcccactct 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 CGGGGTTCAATGGCCACCGTTAATGCAAGCACTAGCTGATCGTTACCCTGCTCCCACTCT 720
```

```
Qy    721 tcgaatcaccggtactggaaatgaccttgatacccttcgtagaacaggtgatcgtttagc 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 TCGAATCACCGGTACTGGAAATGACCTTGATACCCTTCGTAGAACAGGTGATCGTTTAGC 780

Qy    781 taaatttgctcactcattagggttgagatttcaattccatcctctttatatagccaataa 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 TAAATTTGCTCACTCATTAGGGTTGAGATTTCAATTCCATCCTCTTTATATAGCCAATAA 840

Qy    841 taaccacgatcacgatgaagatccttctattatttcctccattgtactactccctgatga 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 TAACCACGATCACGATGAAGATCCTTCTATTATTTCCTCCATTGTACTACTCCCTGATGA 900

Qy    901 aaccctagctatcaactgtgtttttctacctccaccgccttttaaaagaccgcgaaaagtt 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 AACCCTAGCTATCAACTGTGTTTTCTACCTCCACCGCCTTTTAAAAGACCGCGAAAAGTT 960

Qy    961 aaggatttttttgcatagggttaagtcaatgaaccctaaaattgttacaatcgcggagaa 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 AAGGATTTTTTTGCATAGGGTTAAGTCAATGAACCCTAAAATTGTTACAATCGCGGAGAA 1020

Qy   1021 ggaagcaaatcataaccatcctctttttttacaaagattcatcgaggcgttggattatta 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 GGAAGCAAATCATAACCATCCTCTTTTTTTACAAAGATTCATCGAGGCGTTGGATTATTA 1080

Qy   1081 tacagctgtgtttgattcactggaagctacattgccaccgggtagtcgagagaggatgac 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 TACAGCTGTGTTTGATTCACTGGAAGCTACATTGCCACCGGGTAGTCGAGAGAGGATGAC 1140

Qy   1141 agttgaacaagtgtggtttgggagagagattgttgatatcgttgcgatggaaggagataa 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 AGTTGAACAAGTGTGGTTTGGGAGAGAGATTGTTGATATCGTTGCGATGGAAGGAGATAA 1200

Qy   1201 aaggaaagaaagacatgaaaggtttagatcatgggaagttatgttgaggagttgtggatt 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 AAGGAAAGAAAGACATGAAAGGTTTAGATCATGGGAAGTTATGTTGAGGAGTTGTGGATT 1260

Qy   1261 tagtaatgttgctttaagccctttgcattatcacaagctaagcttctttgagacttca 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 TAGTAATGTTGCTTTAAGCCCTTTTGCATTATCACAAGCTAAGCTTCTTTTGAGACTTCA 1320

Qy   1321 ttatccttctgaaggctatcaactcggagtttcgagtaattctttcttcttaggttggca 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 TTATCCTTCTGAAGGCTATCAACTCGGAGTTTCGAGTAATTCTTTCTTCTTAGGTTGGCA 1380

Qy   1381 aaatcaacccctttttctccatctcgtcttggcgttgagaaaaactatcaaatagccaact 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 AAATCAACCCCTTTTCTCCATCTCGTCTTGGCGTTGAGAAAAACTATCAAATAGCCAACT 1440

Qy   1441 tcagagggtaattaagactactgatagtttaggagggatctgaagaaaacgcgtggagtg 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 TCAGAGGGTAATTAAGACTACTGATAGTTTAGGAGGGATCTGAAGAAAACGCGTGGAGTG 1500

Qy   1501 aaaaccctaaataaccagattttctaatgaagttgtagtagtagaaatttgcatggtgaa 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 AAAACCCTAAATAACCAGATTTTCTAATGAAGTTGTAGTAGTAGAAATTTGCATGGTGAA 1560

Qy   1561 gaacaatattgaagaggtattgaaatttcatgtttttttttgttttacttattgatatgaa 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 GAACAATATTGAAGAGGTATTGAAATTTCATGTTTTTTTTTGTTTTACTTATTGATATGAA 1620

Qy   1621 tgttttaaaatttttaacatagaggactaggttgatgatatatagtatttaagttaacta 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 TGTTTTAAAATTTTTAACATAGAGGACTAGGTTGATGATATATAGTATTTAAGTTAACTA 1680

Qy   1681 gtctttgtataacgcaagatcttgatcaacttatttttattttaatta 1729
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 GTCTTTGTATAACGCAAGATCTTGATCAACTTATTTTTATTTTTAATTA 1729

RESULT   3
LES303345/c
LOCUS       LES303345   57047 bp    DNA             PLN       10-APR-2001
DEFINITION  Lycopersicon esculentum lateral suppressor gene, ORF1 and ORF2.
ACCESSION   AJ303345
VERSION     AJ303345.1  GI:13620223
KEYWORDS    lateral suppressor; ls gene; ORF1; ORF2; ORF3; ORF4.
SOURCE      tomato.
  ORGANISM  Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
            Lycopersicon.
REFERENCE   1  (bases 1 to 57047)
  AUTHORS   Rossberg,M., Theres,K., Acarkan,A., Herrero,R., Schmitt,T.,
            Schumacher,K., Schmitz,G. and Schmidt,R.
  TITLE     Comparative sequence analysis reveals extensive microcolinearity in
            the lateral suppressor regions of the tomato, arabidopsis, and
            capsella genomes
  JOURNAL   Plant Cell 13 (4), 979-988 (2001)
  MEDLINE   21178822
REFERENCE   2  (bases 1 to 57047)
  AUTHORS   Theres,K.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JAN-2001) Theres K., Pflanzenzuechtung und
            Ertragsphysiologie, Max-Planck-Institut fuer Zuechtungsforschung,
            Carl-von-Linne-Weg 10, 50829 Koeln, GERMANY
COMMENT     Related sequences AF098674, AJ303342, AJ303343, AJ303344.
FEATURES             Location/Qualifiers
     source          1. .57047
                     /organism="Lycopersicon esculentum"
                     /db_xref="taxon:4081"
                     /chromosome="7"
                     /clone_lib="cosmid library of genomic tomato DNA"
     mRNA            complement(2903. .4189)
                     /gene="ls"
     gene            complement(2903. .4189)
                     /gene="ls"
     CDS             complement(2903. .4189)
                     /gene="ls"
                     /codon_start=1
                     /product="lateral suppressor"
                     /protein_id="CAC36399.1"
                     /db_xref="GI:13620224"
                     /translation="MLGSFGSSSSQSHPHHDEESSDHHQQRRFTATATTITTTTITTS
                     PAIQIRQLLISCAELISQSDFSAAKRLLTILSTNSSPFGDSTERLVHQFTRALSLRLN
                     RYISSTTNHFMTPVETTPTDSSSSSSSLALIQSSYLSLNQVTPFIRFTQLTANQAILEA
                     INGNHQAIHIVDFDINHGVQWPPLMQALADRYPAPTLRITGTGNDLDTLRRTGDRLAK
                     FAHSLGLRFQFHPLYIANNNHDHDEDPSIISSIVLLPDETLAINCVFYLHRLLKDREK
                     LRIFLHRVKSMNPKIVTIAEKEANHNHPLFLQRFIEALDYYTAVFDSLEATLPPGSRE
                     RMTVEQVWFGREIVDIVAMEGDKRKERHERFRSWEVMLRSCGFSNVALSPFALSQAKL
                     LLRLHYPSEGYQLGVSSNSFFLGWQNQPLFSISSWR"
     exon            complement(8920. .10221)
                     /number=4
     mRNA            complement(join(8920. .10221,10358. .10448,11058. .11266,
                     11866. .12087))
     CDS             complement(join(8920. .10221,10358. .10448,11058. .11266,
                     11866. .12087))
                     /note="ORF1"
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="CAC36400.1"
                     /db_xref="GI:13620225"
                     /translation="MAGTAKRHHTGEEGGAGSSDDVFEPSIDKFPVDVISCILEKLDL
                     ESVCSAACVSQTLSSAAAQVFTSLSSLDLSGHYLDEETLEQVARRVQGAKSLTIDCLQ
                     LKNGTSIFNILGEHIEELSLLKCSRVSYHILSAIRERCPNLRSLLIEFAGSEDPQLFE
                     NKLAEMLQKLTLLEVLSIKIRGTYFDVFDIRPLELFLPKSLRKLKLQQTEGDKFVHWL
                     EKIRDIPWFNLQSLSLVLDVISDSLLRTVVNSLPLLVELDLEDRPFMDPTIEDLTNVG
                     LQRVQSCKHLITLAIVRSSMNYRTAFRRVNNMGMFLLSEGCGRLESVKLGGFANVTDA
                     GFSTILNSCRKLKKLEVLNSCLLSDLAFHNMRGVARSLIELRLLSCRLLTSEALEGLS
                     LLSKLEVLDTSGCRSIGNPCLFVISRVTTLTKLNLAEADITDKGLALLGMGNLGITQL
                     CIRGCKRVTDKGIERLFCAEGKIGKTLSLLDVSRMPGITDAAIFTIASAAKALTDLSL
                     RYCFHVTDAGVKMLLDRPNHKVSLLQKLDLYKCRGLSGDWIMSSFCGLRWLGVGGTLL
                     VNKRDDFSTICNVRPWLVVCFDGCEFGCHDGWQFHRIIIHH"
     intron          complement(10222. .10357)
                     /number=3
     exon            complement(10358. .10448)
```

```
                /number=3
intron          complement(10449. .11057)
                /number=2
exon            complement(11058. .11266)
                /number=2
intron          complement(11267. .11865)
                /number=1
exon            complement(11866. .12087)
                /number=1
mRNA            25598. .29176
CDS             25598. .29176
                /note="ORF2"
                /codon_start=1
                /product="hypothetical protein"
                /protein_id="CAC36401.1"
                /db_xref="GI:13620226"
                /translation="MSHKRCHEYFSLNGVLKNVVFMIILLSSSFLVVVSNARKLAEND
                QVGRLLAFKKSSVESDPNGFLNEWTLSSSSPCTWNGISCSNGQVVELNLSSVGLSGLL
                HLTDLMALPTLLRVNFSGNHFYGNLSSIASSCSFEFLDLSANNFSEVLVLEPLLKSCD
                NIKYLNVSGNSIKGVVLKFGPSLLQLDLSSNTISDFGILSYALSNCQNLNLLNFSSNK
                IAGKLKSSISSCKSLSVLDLSRNNLTGELNDLDLGTCQNLTVLNLSFNNLTSVEFPPS
                LANCQSLNTLNIAHNSIRMEIPVELLVKLKSLKRLVLAHNQFFDKIPSELGQSCSTLE
                ELDLSGNRLTGELPSTFKLCSSLFSLNLGNNELSGDFLNTVISSLTNLRYLYLPFNNI
                TGYVPKSLVNCTKLQVLDLSSNAFIGNVPSEFCFAASGFPLETMLLASNYLTGTVPKQ
                LGHCRNLRKIDLSFNNLVGSIPLEIWNLPNLSELVMWANNLTGEIPEGICINGGNLQT
                LILNNNFISGTLPQSISKCTNLVWVSLSSNRLSGEIPQGIGNLANLAILQLGNNSLTG
                PIPRGLGSCRNLIWLDLNSNALTGSIPLELADQAGHVNPGMASGKQFAFVRNEGGTEC
                RGAGGLVEFEGIREERLAILPMVHFCPSTRIYSGRTMYTFTSNGSMIYLDLSYNSLSG
                TIPDNLGSLSFLQVLNLGHNNFTGTIPFNFGGLKIVGVLDLSHNSLQGFIPPSLGGLS
                FLSDLDVSNNNLSGTIPSGGQLTTFPASRYENNSGLCGVPLPPCGSGNGHHSSSIYHH
                GNKKPTTIGMVVGIMVSFICIILLVIALYKIKKTQNEEEKRDKYIDSLPTSGSSSWKL
                STVPEPLSINVATFEKPLRKLTFGHLLEATNGFSSESMIGSGGFGEVYKAQLRDGSTV
                AIKKLVHVTGQGDREFMAEMETIGKIKHRNLVPLLGYCKIGEERLLVYEYMKWGSLES
                VLHDGGKGGMFLDWPARKKIAIGSARGLAFLHHSCIPHIIHRDMKSSNVLLDENFEAR
                VSDFGMARLVNALDTHLSVSTLAGTPGYVPPEYYQSFRCTAKGDVYSYGVILLELLSG
                KRPIDPRVFGDDNNLVGWAKQLHNDKQSHEILDPELITNLSGDAELYHYLKVAFECLD
                EKSYKRPTMIQVMTKFKEVQTDSESDILDGISVKGSILEESQEREP"
exon            39530. .39982
                /note="ORF3"
                /number=1
mRNA            join(39530. .39982,41510. .41698,41775. .42637,42740. .42898,
                42994. .43549)
                /note="ORF3"
CDS             join(39530. .39982,41510. .41698,41775. .42637,42740. .42898,
                42994. .43549)
                /note="ORF3"
                /codon_start=1
                /product="hypothetical protein"
                /protein_id="CAC36402.1"
                /db_xref="GI:13620227"
                /translation="MGGFDDHVAIMGDWMPPSPSPRTFFSSLLGDDGGSRLSFDGTNE
                TSGNLASGPQEHVGNSDGVDTAPAAISEPQPVKLSVPSDQKMNSRGGGLLERIAARAG
                FNAPKLNTESLRPADLRQNQGVRSPYLTIPPGLSPTTLLESPVFLSNSLVQPSPTTGK
                FLFSSGIDNRNSALMMEDTDKRKDNALESINSSSFSFKPVPETAPSLFPGTTSRSWLQ
                VNPSNFSQQGFPNIEVSVHSQNSLVSHRVEATQNPTQNGTLQQSSDFPRFSAEKDVMA
                NNVTLDSRTFQTVGSAVDHSPPLDEPQDEDIDQRGGGDPNVAGAPAEDGYNWRKYGQK
                QVKGSEYPRSYYKCTHPTCPVKKKVERSPEGHITEIIYKGAHNHPKPPPNRRSALGST
                NSLGDLQVDGAEQGASGVNGDLGQANFHKAPGGGGGFDWRNNNRDANLGSEHCNRSAP
                FSAQNNTRLESGDAVDVSSNFSNDEDEDDRGTHGSVSQGYEGEGDESESKRRKLETYS
                ADMTGATRAIREPRVVVQTTSEVDILDDGYRWRKYGQKVVKGNPNPRSYYKCTSAGCN
                VRKHVERASHDLKSVITTYEGKHNHDVPAARNSSHVNSGASNTHPASVTAPAQNHLHR
                PEAAQLQNAMARFDRQPSLGSFGLSGRPPQLGPNPGFSYGMNQQGGLSSLAMAGFHPN
                QNKPGEVPMHPYLGQPRPMHDMGFMFPKEEPKVEPLSDPGPTAYQQFLNSLSLGPQM"
intron          39983. .41508
                /note="ORF3"
                /number=1
exon            41510. .41698
                /note="ORF3"
                /number=2
intron          41699. .41774
                /note="ORF3"
                /number=2
exon            41775. .42637
                /note="ORF3"
                /number=3
intron          42638. .42739
                /note="ORF3"
                /number=3
exon            42740. .42898
                /note="ORF3"
                /number=4
intron          42899. .42993
                /note="ORF3"
                /number=4
exon            42994. .43547
                /note="ORF3"
                /number=5
exon            complement(44881. .44962)
                /note="ORF4"
                /number=9
mRNA            complement(join(44881. .44962,45750. .45982,46133. .46345,
                46506. .46736,48694. .48798,48892. .49346,50526. .50720,
                51517. .51919,52622. .52957))
                /note="ORF4"
CDS             complement(join(44881. .44962,45750. .45982,46133. .46345,
                46506. .46736,48694. .48798,48892. .49346,50526. .50720,
                51517. .51919,52622. .52957))
                /note="ORF4"
                /codon_start=1
                /product="hypothetical protein"
                /protein_id="CAC36403.1"
                /db_xref="GI:13620228"
                /translation="MSGGEYGDHNILLRSTSSASEGDVESQSSPRRTNTRSIKDLLKR
                LDRGFSGRRSSDRDHHHSSSPSPSNRRGVSSTADEILGDSAPPEWAMLLVGCLLGLAT
                GLCVAGFNRGVHVIHEWAWAGTPYDGAAWLRLQRLADTWHRILLIPVLGGVIVGMLHG
                LLGILDQITQSSSTQGQGFDLIAGIFPTVKATQAAITLGTGCSLGPEGPSVDIGKSCA
                YGCSMMMENNRERRIALIAAGAAAGISSGFNAAVAGSFFAIETVLRPLRAENSPPFTT
                AMIILASVISSTVSNAVLGEKQAFTVPTYDMKSAAELPLYLILGMLCGVVSVVFTRLV
                SWFTKGFQFLKEKFGLSDVVCPALGGLGAGVIALRYPGILYWGFTNVDEILHTGKTAS
                APGIGWLAQLVAAKVVATALCKGSGLVGGLYAPSLMIGAAVGAVFGGLAGELINSAIP
                GNAAIAQPQAYALVGMAATLASVCSVPLTSVLLLFELTKDYRILLPLMGAVGLAIWVP
                SVTDQPNEAESSEAKFASKGYSILSPTDENNEGNGSRQSGERNNLELMEVHNSHESFD
                EGLILEDLKVSQAMSNDYLKVSPSQTVKEALECMHDGRQSCVLVVDAEGYLEGILTYG
                DVKRSLFKNHGDSSNKDLSVTDANTCLVSSICTKGISYRGQDCGLLTCYPDTDLAIAK
                QLMEAKGIKQLPVVKRGGEFRRERKRRVIALLHYDSVEETIRREVSHRKSVYQQNEEE
                KDKQIITNGH"
intron          complement(44963. .45749)
                /note="ORF4"
                /number=8
exon            complement(45750. .45982)
                /note="ORF4"
                /number=8
intron          complement(45983. .46132)
                /note="ORF4"
                /number=7

  Query Match           100.0%;  Score 1729;  DB 14;  Length 57047;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 1729;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 cctctgtccttccccccaggtccccttttttttcctttctctctctcctttatttctcttt 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4319 CCTCTGTCCTTCCCCCCAGGTCCCCTTTTTTTCCTTTCTCTCTCTCCTTTATTTCTCTTT 4260

Qy      61 tcataagcatattctttctctctctagggttttcactttcacctgaaatagtgttgttaa 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4259 TCATAAGCATATTCTTTCTCTCTCTAGGGTTTTCACTTTCACCTGAAATAGTGTTGTTAA 4200

Qy     121 attgaatgatatgttaggatcctttggttcttcatcatctcaatctcaccctcatcatga 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4199 ATTGAATGATATGTTAGGATCCTTTGGTTCTTCATCATCTCAATCTCACCCTCATCATGA 4140

Qy     181 tgaagaatcttctgatcatcatcaacagcgtagattcaccgctactgctacaactatcac 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4139 TGAAGAATCTTCTGATCATCATCAACAGCGTAGATTCACCGCTACTGCTACAACTATCAC 4080
```

```
Qy    241 caccaccaccatcactacctcaccagctattcaaatccgccagctactcattagctgtgc 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4079 CACCACCACCATCACTACCTCACCAGCTATTCAAATCCGCCAGCTACTCATTAGCTGTGC 4020

Qy    301 ggagttgatttcgcagtccgatttctcggccgcgaaaagactccttactatattatcaac 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4019 GGAGTTGATTTCGCAGTCCGATTTCTCGGCCGCGAAAAGACTCCTTACTATATTATCAAC 3960

Qy    361 taactcatctccttttggtgattcaactgaacggttagtccatcaatttactcgcgcact 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3959 TAACTCATCTCCTTTTGGTGATTCAACTGAACGGTTAGTCCATCAATTTACTCGCGCACT 3900

Qy    421 ttcccttcgtctcaaccgctatatatcgtcaaccaccaatcatttcatgacacctgttga 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3899 TTCCCTTCGTCTCAACCGCTATATATCGTCAACCACCAATCATTTCATGACACCTGTTGA 3840

Qy    481 aacaactccaactgattcttcttcttcgtcatcattagctctaattcaatcatcatatct 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3839 AACAACTCCAACTGATTCTTCTTCTTCGTCATCATTAGCTCTAATTCAATCATCATATCT 3780

Qy    541 atctctaaaccaagttacccctttcataaggtttactcaattaaccgctaatcaagcgat 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3779 ATCTCTAAACCAAGTTACCCCTTTCATAAGGTTTACTCAATTAACCGCTAATCAAGCGAT 3720

Qy    601 tttagaagcgattaacggtaatcatcaagcaatccacatcgttgatttcgacattaatca 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3719 TTTAGAAGCGATTAACGGTAATCATCAAGCAATCCACATCGTTGATTTCGACATTAATCA 3660

Qy    661 cggggttcaatggccaccgttaatgcaagcactagctgatcgttaccctgctcccactct 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3659 CGGGGTTCAATGGCCACCGTTAATGCAAGCACTAGCTGATCGTTACCCTGCTCCCACTCT 3600

Qy    721 tcgaatcaccggtactggaaatgaccttgataccttcgtagaacaggtgatcgtttagc 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3599 TCGAATCACCGGTACTGGAAATGACCTTGATACCCTTCGTAGAACAGGTGATCGTTTAGC 3540

Qy    781 taaatttgctcactcattagggttgagatttcaattccatcctctttatatagccaataa 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3539 TAAATTTGCTCACTCATTAGGGTTGAGATTTCAATTCCATCCTCTTTATATAGCCAATAA 3480

Qy    841 taaccacgatcacgatgaagatccttctattatttcctccattgtactactccctgatga 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3479 TAACCACGATCACGATGAAGATCCTTCTATTATTTCCTCCATTGTACTACTCCCTGATGA 3420

Qy    901 aaccctagctatcaactgtgttttctacctccaccgccttttaaaagaccgcgaaaagtt 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3419 AACCCTAGCTATCAACTGTGTTTTCTACCTCCACCGCCTTTTAAAAGACCGCGAAAAGTT 3360

Qy    961 aaggatttttttgcatagggttaagtcaatgaaccctaaaattgttacaatcgcggagaa 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3359 AAGGATTTTTTTGCATAGGGTTAAGTCAATGAACCCTAAAATTGTTACAATCGCGGAGAA 3300

Qy   1021 ggaagcaaatcataaccatcctctttttttacaaagattcatcgaggcgttggattatta 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3299 GGAAGCAAATCATAACCATCCTCTTTTTTTACAAAGATTCATCGAGGCGTTGGATTATTA 3240

Qy   1081 tacagctgtgtttgattcactggaagctacattgccaccgggtagtcgagagaggatgac 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3239 TACAGCTGTGTTTGATTCACTGGAAGCTACATTGCCACCGGGTAGTCGAGAGAGGATGAC 3180

Qy   1141 agttgaacaagtgtggtttgggagagagattgttgatatcgttgcgatggaaggagataa 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3179 AGTTGAACAAGTGTGGTTTGGGAGAGAGATTGTTGATATCGTTGCGATGGAAGGAGATAA 3120

Qy   1201 aaggaaagaaagacatgaaaggtttagatcatgggaagttatgttgaggagttgtggatt 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3119 AAGGAAAGAAAGACATGAAAGGTTTAGATCATGGGAAGTTATGTTGAGGAGTTGTGGATT 3060

Qy   1261 tagtaatgttgctttaagccctttgcattatcacaagctaagcttctttgagacttca 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3059 TAGTAATGTTGCTTTAAGCCCTTTTGCATTATCACAAGCTAAGCTTCTTTTGAGACTTCA 3000

Qy   1321 ttatccttctgaaggctatcaactcggagtttcgagtaattctttcttcttaggttggca 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2999 TTATCCTTCTGAAGGCTATCAACTCGGAGTTTCGAGTAATTCTTTCTTCTTAGGTTGGCA 2940

Qy   1381 aaatcaaccccttttctccatctcgtcttggcgttgagaaaaactatcaaatagccaact 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2939 AAATCAACCCCTTTTCTCCATCTCGTCTTGGCGTTGAGAAAAACTATCAAATAGCCAACT 2880

Qy   1441 tcagagggtaattaagactactgatagtttaggagggatctgaagaaaacgcgtggagtg 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2879 TCAGAGGGTAATTAAGACTACTGATAGTTTAGGAGGGATCTGAAGAAAACGCGTGGAGTG 2820

Qy   1501 aaaaccctaaataaccagattttctaatgaagttgtagtagtagaaatttgcatggtgaa 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2819 AAAACCCTAAATAACCAGATTTTCTAATGAAGTTGTAGTAGTAGAAATTTGCATGGTGAA 2760

Qy   1561 gaacaatattgaagaggtattgaaatttcatgttttttttgttttacttattgatatgaa 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2759 GAACAATATTGAAGAGGTATTGAAATTTCATGTTTTTTTTTGTTTTACTTATTGATATGAA 2700

Qy   1621 tgttttaaaatttttaacatagaggactaggttgatgatatatagtatttaagttaacta 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2699 TGTTTTAAAATTTTTAACATAGAGGACTAGGTTGATGATATATAGTATTTAAGTTAACTA 2640

Qy   1681 gtctttgtataacgcaagatcttgatcaacttatttttatttttaatta 1729
          |||||||||||||||||||||||||||||||||||||||||||||||
Db   2639 GTCTTTGTATAACGCAAGATCTTGATCAACTTATTTTTATTTTTAATTA 2591
```

RESULT 4
A84080
LOCUS A84080 1296 bp DNA PAT 21-JAN-2000
DEFINITION Sequence 9 from Patent WO9846759.
ACCESSION A84080
VERSION A84080.1 GI:6733220
KEYWORDS .
SOURCE potato.
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 1296)
AUTHORS Theres,N.
TITLE PLANTS WITH CONTROLLED SIDE-SHOOT FORMATION AND/OR CONTROLLED ABSCISSION AREA FORMATION
JOURNAL Patent: WO 9846759-A 9 22-OCT-1998;
THERES NIKOLAUS (DE)
FEATURES Location/Qualifiers
source 1. .1296
/organism="Solanum tuberosum"
/db_xref="taxon:4113"
BASE COUNT 354 a 303 c 229 g 410 t
ORIGIN

Query Match 70.7%; Score 1223.2; DB 9; Length 1296;
Best Local Similarity 97.1%; Pred. No. 3.2e-251;
Matches 1259; Conservative 0; Mismatches 28; Indels 9; Gaps 1;

```
Qy    131 atgttaggatcctttggttcttcatcatctcaatctcaccctcatcatgatgaagaatct 190
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGTTAGGATCCTTTGGTTCTTCATCATCTCAATCTCACCCTCATCATGATGAAGAATCT 60

Qy    191 tctgatcatcatcaacagcgtagattcaccgctactgctacaactatcaccaccaccacc 250
          |||||||||||||||| |||||||||||||||||| ||||||||||||||||||||||||
Db     61 TCTGATCATCATCAACGGCGTAGATTCACCGCTACTACTACAACTATCACCACCACCACC 120

Qy    251 atcactacctcaccagctattcaaatccgccagctactcattagctgtgcggagttgatt 310
          |  || ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 ACAACGACCTCACCAGCTATTCAAATCCGCCAGCTACTCATTAGCTGTGCGGAGTTGATT 180

Qy    311 tcgcagtccgatttctcggccgcgaaaagactccttactatattatcaactaactcatct 370
          |||| |||||||||||||||||||||||||||||||| |||||||||||||||||| |||
Db    181 TCGCGGTCCGATTTCTCGGCCGCGAAAAGACTCCTTACCATATTATCAACTAACTCTTCT 240
```

```
Qy    371 ccttttggtgattcaactgaacggttagtccatcaatttactcgcgcactttcccttcgt 430
          ||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db    241 CCTTTTGGTGATTCAACTGAACGGTTAGTCCATCAGTTTACTCGCGCACTTTCCCTTCGT 300

Qy    431 ctcaaccgctatatatcgtcaaccaccaatcatttcatgacacctgttgaaacaactcca 490
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CTCAACCGCTATATATCGTCAACCACCAATCATTTCATGACACCTGTTGAAACAACTCCA 360

Qy    491 actgattcttcttct---------tcgtcatcattagctctaattcaatcatcatatcta 541
          |||||||||| |||         |||||||||||||||||||||||||||||||||||
Db    361 ACTGATTCTTCATCTTCGTTGCCATCGTCATCATTAGCTCTAATTCAATCATCATATCAT 420

Qy    542 tctctaaaccaagttacccctttcataaggtttactcaattaaccgctaatcaagcgatt 601
          |||||||| |||||||||||||| ||||||||||||||||||||||||||||||||||||
Db    421 TCTCTAAATCAAGTTACCCCTTTTATAAGGTTTACTCAATTAACCGCTAATCAAGCGATT 480

Qy    602 ttagaagcgattaacggtaatcatcaagcaatccacatcgttgatttcgacattaatcac 661
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TTAGAAGCGATTAACGGTAATCATCAAGCAATCCACATCGTTGATTTCGACATTAATCAC 540

Qy    662 ggggttcaatggccaccgttaatgcaagcactagctgatcgttaccctgctcccactctt 721
          |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db    541 GGGGTTCAATGGCCACCGTTAATGCAAGCACTAGCTGATCGTTACCCTGCTCCTACTCTT 600

Qy    722 cgaatcaccggtactggaaatgaccttgataccttcgtagaacaggtgatcgtttagct 781
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 CGAATCACCGGTACTGGAAATGACCTTGATACCCTTCGTAGAACAGGTGATCGTTTAGCT 660

Qy    782 aaatttgctcactcattagggttgagatttcaattccatcctctttatatagccaataat 841
          |||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db    661 AAATTTGCTCACTCATTAGGGTTGAGATTTCAATTCCATCCTCTTTATATCGCCAATAAT 720

Qy    842 aaccacgatcacgatgaagatccttctattatttcctccattgtactactccctgatgaa 901
          |||| |||||||| ||||||||||||||||||||||||||||||| ||||||||||||
Db    721 AACCGCGATCACGGTGAAGATCCTTCTATTATTTCCTCCATTGTACTTCTCCCTGATGAA 780

Qy    902 accctagctatcaactgtgttttctacctccaccgccttttaaaagaccgcgaaaagtta 961
          |||||||||||||||||||||||| |||||||||||||||||||||||| |||
Db    781 ACCCTAGCTATCAACTGTGTTTTCTATCTCCACCGCCTTTTAAAAGACCGCGAAAAATTA 840

Qy    962 aggatttttttgcatagggttaagtcaatgaaccctaaaattgttacaatcgcggagaag 1021
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 AGGATTTTTTTGCATAGGGTTAAGTCAATGAACCCTAAAATTGTTACAATCGCGGAGAAG 900

Qy   1022 gaagcaaatcataaccatcctctttttttacaaagattcatcgaggcgttggattattat 1081
          |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db    901 GAAGCAAATCATAACCATCCTCTTTTTTTACAAAGATTTATCGAGGCGTTGGATTATTAT 960

Qy   1082 acagctgtgtttgattcactggaagctacattgccaccgggtagtcgagagaggatgaca 1141
          |||||||||||||||| |||||||||||||||||||||||| ||||||||||||
Db    961 ACAGCTGTGTTTGATTCATTGGAAGCTACATTGCCACCGGGTAGTCGTGAGAGGATGACA 1020

Qy   1142 gttgaacaagtgtggtttgggagagagattgttgatatcgttgcgatggaaggagataaa 1201
          |||||||||||||||||||||||| |||||||||||||| ||||||||||||||||||
Db   1021 GTTGAACAAGTGTGGTTTGGGAGAGAAATTGTTGATATCGTGGCGATGGAAGGAGATAAA 1080

Qy   1202 aggaaagaaagacatgaaaggtttagatcatgggaagttatgttgaggagttgtggattt 1261
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 AGGAAAGAAAGACATGAAAGGTTTAGATCATGGGAAGTTATGTTGAGGAGTTGTGGATTT 1140

Qy   1262 agtaatgttgctttaagccctttttgcattatcacaagctaagcttcttttgagacttcat 1321
          |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db   1141 AGTAATGTTGCTTTAAGCCCTTTTGCATTATCACAAGCTAAGCTTCTTTTGAGACTACAT 1200

Qy   1322 tatccttctgaaggctatcaactcggagtttcgagtaattcttttcttcttaggttggcaa 1381
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 TATCCTTCTGAAGGCTATCAACTCGGAGTTTCGAGTAATTCTTTCTTCTTAGGTTGGCAA 1260

Qy   1382 aatcaaccccttttctccatctcgtcttggcgttga 1417
          |||||||| |||||||||||||||||||||||||
Db   1261 AATCAACCTCTTTTCTCCATCTCGTCTTGGCGTTGA 1296
```

```
RESULT   5
AC005223/c
LOCUS       AC005223    87967 bp    DNA             PLN       30-JAN-1999
DEFINITION  Arabidopsis thaliana chromosome I BAC T5A14 genomic sequence,
            complete sequence.
ACCESSION   AC005223
VERSION     AC005223.1  GI:4096078
KEYWORDS    HTG.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
            Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
            Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 87967)
  AUTHORS   Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,
            Altafi,H., Araujo,R., Huizar,L., Rowley,D., Buehler,E., Dunn,P.,
            Gonzalez,A., Kremenetskaia,I., Kim,C., Lenz,C., Li,J., Liu,S.,
            Luros,S., Schwartz,J., Shinn,P., Toriumi,M., Vysotskaia,V.S.,
            Walker,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 87967)
  AUTHORS   Federspiel,N.A., Palm,C.J., Conway,A.B., Kurtz,D.B., Conway,A.R.,
            Au,M., Araujo,R., Buehler,E., Dewar,K., Feng,J., Kim,C., Li,Y.,
            Oji,O., Osborne,B.I., Shinn,P., Sun,H., Toriumi,M., Vyotskaia,V.,
            Yu,G., Ecker,J., Theologis,A. and Davis,R.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-JUL-1998) DNA Sequencing and Technology Center,
            Stanford University, 855 California Avenue, Palo Alto, CA 94304,
            USA
REFERENCE   3  (bases 1 to 87967)
  AUTHORS   Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,
            Altafi,H., Araujo,R., Huizar,L., Rowley,D., Buehler,E., Dunn,P.,
            Gonzalez,A., Kremenetskaia,I., Kim,C., Lenz,C., Li,J., Liu,S.,
            Luros,S., Schwartz,J., Shinn,P., Toriumi,M., Vyotskaia,V.,
            Walker,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-JAN-1999) DNA Sequencing and Technology Center,
            Stanford University, 855 California Avenue, Palo Alto, CA 94304,
            USA
REFERENCE   4  (bases 1 to 87967)
  AUTHORS   Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,
            Altafi,H., Araujo,R., Huizar,L., Rowley,D., Buehler,E., Dunn,P.,
            Gonzalez,A., Kremenetskaia,I., Kim,C., Lenz,C., Li,J., Liu,S.,
            Luros,S., Schwartz,J., Shinn,P., Toriumi,M., Vyotskaia,V.,
            Walker,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-JAN-1999) DNA Sequencing and Technology Center,
            Stanford University, 855 California Avenue, Palo Alto, CA 94304,
            USA
COMMENT     On Jan 5, 1999 this sequence version replaced gi:3779015.
            Bases 1-9066 of clone of clone T5A14 overlap with bases
            100,099-109,160 of 'IGF' clone F20N2, gb|AC002328.
            e-mail for correspondence: arab@sequence.stanford.edu
            Genes with similarity to proteins in the databases are described as
            'putative', '-like' or 'similar to'.  Genes that have EST
            similarity but no significant protein similarity are described as
            'unknown proteins'.  Genes that are annotated based only on gene
            prediction software are described as 'hypothetical proteins'.
            The software programs used to predict genes include:  Grail
            (Informatics Group, Oak Ridge National Laboratory,
            http://compbio.ornl.gov/section/index.html), GENSCAN (Chris Burge,
            http://gnomic.stanford.edu/~chris/GENSCANW.html), Fexa (V.Solovyev
            & A.Salamov, Sanger Centre,  http://genomic.sanger.ac.uk/), and
            NetPlantGene (S.M. Hebsgaard, et al., CBS, Technical University of
            Denmark,http://www.cbs.dtu.dk/NetPlantGene.html).
FEATURES              Location/Qualifiers
     source           1. .87967
                      /organism="Arabidopsis thaliana"
                      /db_xref="taxon:3702"
                      /chromosome="I"
     misc_feature     1. .9066
                      /note="overlap with bases 100,099-109,160 of 'IGF' clone
```

```
F20N2, gb|AC002328.  This region is annotated in the F20N2
entry, gb|AC002328.
"
gene            complement(11165. .13054)
                /gene="T5A14.1"
CDS             complement(join(11165. .11468,11557. .12666,12801. .13054))
                /gene="T5A14.1"
                /note="putative pollen specific protein; nearly identical
                to pollen specific proteins from tobacco, gi|128592 and
                Brassica napus, gi|99809 and very similar to tomato
                pectinesterase gi|1944575"
                /codon_start=1
                /protein_id="AAD10638.1"
                /db_xref="GI:4204257"
                /translation="MKGGVKLLAVCLCVATATVMMVQAEDPYFHHVWNVTYGTASPLG
                VPQQVILINGQFPGPNINSTSNNNVIVNVFNNLDEPFLITWAGIQHRKNCWQDGTAGT
                MCPIPPGQNFTYHFQPKDQIGSYFYYPTTAMHRAAGGFGGLRVNSRLLIPVPYADPED
                DYTILINDWYTKSHTQLKKFLDSGRTIGRPDGILINGKSGKTDGSDKPLFTLKPGKTY
                RVRICNVGLKASLNFRIQNHKMKLVEMEGSHVLQNDYDSLDVHVGQCFGVIVTADQEP
                KDYYMIASTRFLKKPLTTTGLLRYEGGKGPASSQLPAAPVGWAWSLNQYRSFRWNLTA
                SAARPNPQGSYHYGKINITRTIKLVNTQGKVDGKLRYALSGVSHTDPETPLKLAEYFG
                VADKVFKYDTISDNPNPDQIKNIKIEPNVLNITHRTFIEVVFENHERSVQSWHLDGYS
                FFAVAVEPGTWTPEKRKNYNLLDAVSRHTVQVYPKCWAAILLTFDNCGMWNIRSENAE
                RRYLGQQLYASVLSPEKSLRDEYNMPETSLQCGLVKGKPKVNPYAGA"
gene            14409. .16462
                /gene="T5A14.2"
CDS             join(14409. .14653,14745. .14874,15074. .16077,16165. .16462)
                /gene="T5A14.2"
                /note="putative pollen specific protein; highly similar to
                pollen specific proteins from tobacco, gi|128592 and
                Brassica napus, gi|99809 and very similar to tomato
                pectinesterase gi|1944575"
                /codon_start=1
                /protein_id="AAD10639.1"
                /db_xref="GI:4204258"
                /translation="MEGRLLTVLVCLVSTVAIVNAGDPYFFHTWNVTYGTASPLGVPQ
                KVILINGQFPGPNLNSTSNNNVVINVFNHLDEPFLLTCLSQSIICAQAIVGPRLIFYR
                TFPIRSSFDGIQGKGFSADRLSRDQIGSYFYYPTTSLHRFAGGFGGLRVNSRLLIPVP
                YADPEDDYTVLLGDWYTAGHTALKNFLDSGRTLGLPNGVLINGKSGKVGGKNEPLFTM
                KPGKTYKYRLCNVGFKSTLNFRIQNHKMKLVEMEGSHVIQNDYDSLDVHVGQCFSVLV
                TANQAAKDYYMVASTRFLKKELSTVGVIRYEGSNVQASTELPKAPVGWAWSLNQFRSF
                RWNLTSNAARPNPQGSYHYGKINITRSIKLVNSKSVVDGKVRFGFNGVSHVDTETPLK
                LAEYFQMSEKVFKYNVIKDEPAAKITALTVQPNVLNITFRTFVEIIFENHEKTMQSFH
                LDGYSFFAVASEPGRWTPEKRENYNLLDAVSRHTVQVYPKSWSAILLTFDNAGMWNIR
                SENLERKYLGEQLYVSVLSPEKSLRDEYNIPLNTNLCGIVKGLPLPAHYS"
gene            complement(18074. .22021)
                /gene="T5A14.3"
CDS             complement(join(18074. .19480,19572. .19699,19771. .19808,
                19888. .20103,20189. .20350,20454. .20632,20721. .20827,
                20920. .21078,21166. .21249,21838. .22021))
                /gene="T5A14.3"
                /note="Similar to Kinesin proteins; Contains kinesin motor
                domain protein motif and kinesin heavy chain signature
                motif"
                /codon_start=1
                /protein_id="AAD10640.1"
                /db_xref="GI:4204259"
                /translation="MERTRSKPVRNLPETIHSLLGLKSHMTSDWVKSVCNIAKNTSST
                SKKEEDDFVSVDLQSIRDQLSALTVQVNDQNKLRRQILNEFLDLKGNIRVFCRVKPLG
                ATEKLRPPVASDTRNVIIKLSETKRKTYNFDRVFQPDSSQDDVFLEIEPVIKSVIDGY
                NACIFAYGQTGTGKTYTMEGLPNSPGIVPRAIKGLFKQVEESNHMFTIHFSMLEIYMG
                NLKDLLLSEATKPISPIPPSLSIHTDPNGEIDIENLVKLKVDDFNEILRLYKVGCRSR
                ATASTNSNSVSSRSHCMIRVSVTSLGAPERRRETNKIWLVDLGGSERVLKTRATGRRF
                DEGKAINLSLSALGDVINSLQRKNSHIPYRNSKLTQVLKDSLGQDSKTLMLVHISPKE
                DDLCETICSLNFATRAKNIHLGQDESTEEQAKKEAVMMNLQKMMEKIEQEREMSLRKM
                RNLNETLEKLTGKPHVIEEEEKDVVREVIHVTPKKPRNKSRRASDVFPSFMRPTASSN
                RRLSGADFSVTPNSSSFKSRRNSMISVRAESACLPVKKKKNRFDSACDSSDRSVSKST
                SIMRQNTADDATVYSQDISECDIKLVVSEHKPKPLQMGPGSATKSRSNISNFEKDVMQ
                KIGGTEFSRINSWLRSQSENRSYVLDKTQLPATHFLENLNRSLEKSPTQSFTTEKITG
                NELEGIEETKTNETVVNPTLMLKKLFELQCLCSAEEEDQILSRFPIPGYEDDDESRYP
                PILENDGFSQHIDNEWFGVNNYSADWERDSPATIPLLECEPDLKQLLPELGLDRSLKP
                RGLAVAEGAAPPLLRAQETLGERGKRLCSGVKCFYKRKRFSKILSFSVVVAGKGPTFM
                QKLQALCFRILLGLGFMDVGFGNDFFNGLTK"
gene            complement(25568. .26391)
                /gene="T5A14.4"
CDS             complement(join(25568. .25615,25969. .26031,26184. .26243,
                26332. .26391))
                /gene="T5A14.4"
                /note="Hypothetical protein"
                /codon_start=1
                /protein_id="AAD10641.1"
                /db_xref="GI:4204260"
                /translation="MSAKDLVDSVFTLGGWSNFLCSEKKLQQPVDVALVFIGRELLSS
                DVSSNQNSDPVLVNTLKSLNSEDPYVLQVATH"
gene            27887. .36177
                /gene="T5A14.5"
CDS             join(27887. .28076,28158. .28381,28449. .28617,28726. .28872,
                29142. .29263,29411. .29523,29938. .30077,30163. .30328,
                30411. .31353,31534. .31686,31842. .31899,32440. .32573,
                32663. .32787,32908. .33250,33402. .33585,33706. .34323,
                34483. .35940,36050. .36177)
                /gene="T5A14.5"
                /note="Hypothetical protein"
                /codon_start=1
                /protein_id="AAD10642.1"
                /db_xref="GI:4204261"
                /translation="MSRVEIEEDTEGDRISTNDYYFERIGEPISIKEDDAQYDLENPP
                SQPLAISERHAVLFVAHSSGFFVGRTNDVISASKNSNGNGDKVFIQDLSLVDVPVGDV
                RILSLSADDSILAVTVAADIHFFSVDSLLKKVHVSEDAKPSFSYSPDESGFVKDFRWR
                RNDKHSYLVLSNTGKLFHGIDNAPPRHVMDAVDAVEWSSKGSYIAVAQDNSLRIFSSK
                FNEKRCIALSFDSWIGDSDEDCFVKVDSIRWVRNNCILLGCFQLIEGREENYLVQVIR
                SPDGKISDGSTNLVALSFSDLFPCSMDDLVPVGVGPHLLFSYIDQCKLAVTANRKSID
                EHIVLLDWSSGDDKSAVSVVDIDRETFLPRIGLQENNDDNTVMGLCIDRVSIEGTVNV
                RSGDDELKELQPYFVLVCLTLEGKLVMFNVASVAGRPASSDTDLASSSDIEDAYTPLI
                EDDLSKQSSEKHQQLNIAVQNDQKHLNTEKFSTEQRLPNENIFSKEPESVKSSVSGDN
                NKKQEPYAEKPLQVEDAQQSMIPRLSGTSFGQLPMSLGYDTNKFAGFGPALPVSEKLQ
                KDIFAQSNSMHLQANVESKSTAAFFGSPGLQNAILQSPQNTSSQPWSSGKSVSPPDFV
                SGPFPSMRDTQHKQSVQSGTGYVNPPMSIKDKSVQVIETGRVSALSNLSPLLGQNQDT
                NEGVEKIEPIPSIRASQLSQQVKSSFEKSASHQQHKTPLSTGPLRLEHNMSNQPSNIN
                EMAREMDTLLQSIEGPGGFKDSCAFILKSNVEELEQGLESLAGKCQTWKSTIHEQQAE
                IQHLLDKTIQVLAKKTYMEGMYKQTADNQYWQLWNRQKLNPELEAKRQHIMKLNKDLT
                HQLIELERYFNRLELDRYNEDGGHPVARRGVPNRSAPSRRVQSLHSLHNTMSSQLAAA
                EQLSECLSKQMTYLKIDSPVKKNVKQELFETIGIPYDASFSSPDAVKAKNASSAKNLL
                LSSIPASINQQSRQRQSSAMKNSDPETARRRRESLDRNWAAFEPPKTTVKRMLLQEQQ
                KTGMNQQTVLSERLRSANNTQDRSLLHVKDHASPVVSSNKGIMESFQQDTSEAQSTPF
                KTRPPMPQSNSPFTISPISASKPSFNWSGNKSSNTTSYAEESAPSQIKDTRTVSQPGG
                SSFLPKRPVASTVLEQTEKKAGEFKFSEAKANAFVETAAGSVQRLSTTSSGSDFESSK
                GFGAQFSTMSSGAPASSFSSKSLFGFNSSSSIPGDKFTFPAVTAPLSGTPLDSTSTLF
                TASSAPVSSSSQDPVTSTSTVSATGFNVPFGKPLTSVKVDLNQAAPSTPSPSPGPTAG
                FTFNLPALSPSSPEMVSSSTGQSSLFPPSAPTSQVSSDQASATSSLTDSSRLFSSTSL
                SSTPPITPPDAFQSPQVSTPSSAVPITEPVSEPKKPEAQSSSILSTQSTVDSVANATK
                TQNEPLPVKSEISNPGTTVTPVSSSGFLSGFSSGTQSSLASMAAPSFSWPGSSQPQQL
                SSTPAPFPASSPTSASPFGEKKDIVDTQEDEMDEEAPEASQTTELSMGSFGGFGLGST
                PNPGAPKTNPFGGPFGNATTTTSNPFNMTVPSGELFKPASFNFQNPQPSQPAGFGSFS
                VTPSQTPAQSGFGQPSQIGGGQQALGSVLGSFGQSRQIGAGLPGATFGSPTGFGGSNP
                GSGLPNAPASGGFAAAGSSATGGFAAMASAGRGFAGASSTPTGGFAALASGSGGFAGA
                APGGGGGGFGGLGSGTGGFGGFAPQGSSGGFAGAAGGGGFGGFGGQAQGQAGGGGFSA
                FGGNSGATGKPSELFTQMRK"
gene            37783. .39494
                /gene="T5A14.6"
```

```
Query Match             17.0%;  Score 294.2;  DB 12;  Length 87967;
Best Local Similarity   56.3%;  Pred. No. 1.1e-52;
Matches  719;  Conservative    0;  Mismatches  483;  Indels   75;  Gaps    6;

Qy    222 ctactgctacaactatcaccaccaccaccatcactacctcaccagctattcaaatccgcc 281
          || || ||      ||| || || || |   | | |||||  |  |   |||  |||| |
Db   6763 CTCCTCCTTTGTGCATCGCCTCCTCCTCGGCCGCAACCTCCGCCTCACATCACCTCCGTC 6704

Qy    282 agctactcattagctgtgcggagttgatttcgcagtccgatttctcggccgcgaaaagac 341
            || ||  | | |  |||| | ||  | || |||||  | ||| | |||||  |||
Db   6703 GTCTTCTTTTCACCGCTGCGAATTTCGTCTCCCAGTCAAACTTCACCGCCGCTCAAAACT 6644

Qy    342 tccttactatattatcaactaactcatctccttttggtgattcaactgaacggttagtcc 401
          | ||  | ||  | ||   |||||| ||||||   || || || || || ||  | || |
Db   6643 TACTCTCAATCCTCTCCCTTAACTCTTCTCCTCACGGCGACTCCACCGAGCGACTTGTAC 6584
```

```
Qy    402 atcaatttactcgcgcactttcccttcgtctcaaccgcta---------- 441
Db   6583 ACCTCTTCACTAAAGCCTTGTCCGTACGAATCAACCGTCAGCAACAAGATCAGACGGCTG 6524

Qy    442 -------tatatcgtcaaccaccaatcatttcatgacacctgttgaaacaac-------- 486
Db   6523 AAACGGTTGCCACGTGGACGACGAACGAAATGACGATGAGTAACTCCACGGTGTTCACGA 6464

Qy    487 --------tccaactgattcttcttcttcgtcatcattagctctaattcaatcatc---- 534
Db   6463 GCAGTGTATGCAAAGAACAGTTCTTGTTTCGAACCAAGAACAACAATTCTGACTTCGAGT 6404

Qy    535 -------atatctatctctaaaccaagttacccctttcataaggtttactcaattaaccg 587
Db   6403 CTTGTTACTATCTTTGGCTAAACCAACTAACGCCGTTTATTCGGTTCGGTCATTTAACGG 6344

Qy    588 ctaatcaagcgatttagaagcgattaacggtaatcatca---agcaatccacatcgttg 644
Db   6343 CGAACCAAGCTATCCTCGACGCGACGGAGACAAACGATAACGGAGCTCTACATATACTTG 6284

Qy    645 atttcgacattaatcacggggttcaatggccaccgttaatgcaagcactagctgatcgtt 704
Db   6283 ATTTAGATATATCACAAGGACTTCAATGGCCTCCATTGATGCAAGCCCTAGCAGAGAGGT 6224

Qy    705 ac---------------cctgctcccactcttcgaatcaccggtactggaaatgaccttg 749
Db   6223 CATCAAACCCTAGCAGTCCACCTCCATCTCTCCGCATAACCGGATGCGGTCGAGATGTAA 6164

Qy    750 ataccccttcgtagaacaggtgatcgtttagctaaatttgctcactcattagggttgagat 809
Db   6163 CCGGATTAAACCGAACTGGAGACCGGTTAACCCGGTTCGCTGACTCTTTAGGTCTCCAAT 6104

Qy    810 ttcaattccatcctctttatatagccaataataaccacgatcacgatgaagatccttcta 869
Db   6103 TCCAGTTTCACACGCTAGTGATCGTAGAAGAAGATCTCGCCGGACTTTTGCTACAGATCC 6044

Qy    870 ttatttcctccattgtactactccctgatgaaaccctagctatcaactgtgttttctacc 929
Db   6043 GATTGTTAGCTCTCTCAGCCGTACAAGGAGAGACCATTGCCGTCAATTGTGTTCACTTCC 5984

Qy    930 tccaccgccttttaaaagac---cgcgaaaagttaaggatttttttgcatagggttaagt 986
Db   5983 TCCACAAAATATTTAACGACGATGGAGATATGATCGGTCACTTCTTGTCAGCGATCAAGA 5924

Qy    987 caatgaaccctaaaattgttacaatcgcggagaaggaagcaaatcataaccatcctcttt 1046
Db   5923 GCTTAAACTCTAGAATCGTTACAATGGCAGAGAGAGAAGCTAATCATGGAGATCACTCGT 5864

Qy   1047 ttttacaaagattcatcgaggcgttggattattatacagctgtgtttgattcactggaag 1106
Db   5863 TCTTGAATAGATTCTCTGAGGCAGTGGATCATTACATGGCGATCTTTGATTCGTTGGAAG 5804

Qy   1107 ctacattgccaccgggtagtcgagagaggatgacagttgaacaagtgtggtttgggagag 1166
Db   5803 CGACGTTGCCGCCAAATAGCCGAGAGAGACTAACCCTAGAGCAACGGTGGTTCGGTAAGG 5744

Qy   1167 agattgttgatatcgttgcgatggaaggagataaaaggaaagaaagaacatgaaaggttta 1226
Db   5743 AGATTTTGGATGTTGTGGCGGCGGAAGAGACGGAGAGAAAGCAAAGACATCGGAGGTTTG 5684

Qy   1227 gatcatgggaagttatgttgaggagttgtggatttagtaatgttgctttaagccctttg 1286
Db   5683 AGATTTGGGAAGAGATGATGAAGAGGTTTGGTTTCGTTAACGTTCCTATTGGAAGCTTTG 5624

Qy   1287 cattatcacaagctaagcttcttttgagacttcattatccttctgaaggctatcaactcg 1346
Db   5623 CTTTGTCTCAAGCTAAGCTTCTTCTTAGACTTCATTATCCTTCAGAAGGTTATAATCTTC 5564

Qy   1347 gagtttcgagtaattctttcttcttaggttggcaaaatcaacccctttttctccatctcgt 1406
Db   5563 AGTTCCTTAACAATTCTTTGTTTCTTGGCTGGCAAAATCGTCCCCTCTTCTCCGTTTCGT 5504

Qy   1407 cttggcgttgagaaaaa 1423
Db   5503 CGTGGAAATGATCCAAA 5487
```

```
RESULT   6
AC002328
LOCUS       AC002328   109171 bp    DNA             PLN       28-JUN-2000
DEFINITION  Genomic sequence for Arabidopsis thaliana BAC F20N2 from chromosome
            I, complete sequence.
ACCESSION   AC002328
VERSION     AC002328.2  GI:7798720
KEYWORDS    HTG:
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
            Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
            Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 109171)
  AUTHORS   Khan,S., Brooks,S., Buehler,E., Chao,Q., Johnson-Hopson,C., Kim,C.,
            Shinn,P., Altafi,H., Bei,Q., Chin,C., Chiou,J., Choi,E., Conn,L.,
            Conway,A., Gonzales,A., Hansen,N., Howng,B., Koo,T., Lam,B.,
            Lee,J., Lenz,C., Li,J., Liu,A., Liu,K., Liu,S., Mukharsky,N.,
            Nguyen,M., Palm,C., Pham,P., Sakano,H., Schwartz,J., Southwick,A.,
            Thaveri,A., Toriumi,M., Vaysberg,M., Yu,G., Federspiel,N.A.,
            Theologis,A. and Ecker,J.R.
  TITLE     Genomic sequence for Arabidopsis thaliana BAC F20N2 from chromosome
            I
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 109171)
  AUTHORS   Ecker,J.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-JUL-1997) Arabidopsis thaliana Genome Center,
            Department of Biology, University of Pennsylvania, 38th Street and
            Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA
REFERENCE   3  (bases 1 to 109171)
  AUTHORS   Ecker,J.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-AUG-1998) Arabidopsis thaliana Genome Center,
            Department of Biology, University of Pennsylvania, 38th Street and
            Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA
REFERENCE   4  (bases 1 to 109171)
  AUTHORS   Shinn,P.P., Buehler,E.E., Dunn,P.P., Feng,J.J., Kim,C.C., Li,Y.Y.,
            Walker,M.M., Altafi,H.H., Araujo,R.R., Conn,L.L., Conway,A.A.B.,
            Gonzalez,A.A., Hansen,N.N.F., Huizar,L.L., Kremenetskaia,I.I.,
            Lenz,C.C., Li,J.J., Liu,S.S., Luros,S.S., Rowley,D.D.,
            Schwartz,J.J., Toriumi,M.M., Vysotskaia,V.V., Yu,G.G.,
            Davis,R.R.W., Federspiel,N.N.A., Theologis,A.A. and Ecker,J.J.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-DEC-1998) Arabidopsis thaliana Genome Center,
            Department of Biology, University of Pennsylvania, 38th Street and
            Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA
REFERENCE   5  (bases 1 to 109171)
  AUTHORS   Ecker,J.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAY-2000) Arabidopsis thaliana Genome Center,
            Department of Biology, University of Pennsylvania, 38th Street and
            Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA
REFERENCE   6  (bases 1 to 109171)
  AUTHORS   Cheuk,R., Shinn,P., Brooks,S., Buehler,E., Chao,Q.,
            Johnson-Hopson,C., Khan,S., Kim,C., Altafi,H., Bei,B., Chin,C.,
            Chiou,J., Choi,E., Conn,L., Conway,A., Gonzalez,A., Hansen,N.,
            Howing,B., Koo,T., Lam,B., Lee,J., Lenz,C., Li,J., Liu,A., Liu,J.,
            Liu,S., Mukharsky,N., Nguyen,M., Palm,C., Pham,P., Sakano,H.,
            Schwartz,J., Southwick,A., Thaveri,A., Toriumi,M., Vaysberg,M.,
            Yu,G., Davis,R., Federspiel,N., Theologis,A. and Ecker,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-JUN-2000) Arabidopsis thaliana Genome Center,
            Department of Biology, University of Pennsylvania, 38th and
            Hamilton Walk, Philadelphia, PA 19104-6018, USA
COMMENT     On May 13, 2000 this sequence version replaced gi:3492855.
FEATURES             Location/Qualifiers
     source          1. .109171
```

/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/chromosome="1"
/clone="F20N2"
CDS 2237. .3574
/note="hypothetical protein"
/codon_start=1
/evidence=not_experimental
/product="F20N2.1"
/protein_id="AAF79493.1"
/db_xref="GI:8778485"
/translation="MLTSFKSSSSSSEDATATTTENPPPLCIASSSAATSASHHLRRL
LFTAANFVSQSNFTAAQNLLSILSLNSSPHGDSTERLVHLFTKALSVRINRQQQDQTA
ETVATWTTNEMTMSNSTVFTSSVCKEQFLFRTKNNNSDFESCYYLWLNQLTPFIRFGH
LTANQAILDATETNDNGALHILDLDISQGLQWPPLMQALAERSSNPSSPPPSLRITGC
GRDVTGLNRTGDRLTRFADSLGLQFQFHTLVIVEEDLAGLLLQIRLLALSAVQGETIA
VNCVHFLHKIFNDDGDMIGHFLSAIKSLNSRIVTMAEREANHGDHSFLNRFSEAVDHY
MAIFDSLEATLPPNSRERLTLEQRWFGKEILDVVAAEETERKQRHRRFEIWEEMMKRF
GFVNVPIGSFALSQAKLLLRLHYPSEGYNLQFLNNSLFLGWQNRPLFSVSSWK"
CDS complement(join(7611. .8987,9078. .9168,9453. .9658,
9742. .9891))
/note="similar to uridine kinase sp|032033|URK_BACSU"
/codon_start=1
/evidence=not_experimental
/product="F20N2.2"
/protein_id="AAF79512.1"
/db_xref="GI:8778504"
/translation="MEEVTRSVLAEEILKRLDLENLCSVACVSTTLRSAVVSGVLPSL
TSLDLSVFSPDDETLNHVLRGCIGLSSLTLNCLRLNAASVRGVLGPHLRELHLLRCSL
LSSTVLTYIGTLCPNLRVLTLEMADLDSPDVFQSNLTQMLNGCPYLESLQLNIRGILV
DATAFQSVRFSLPETLKALRLQPLLESEAILLMNRFKVTGTYLSQPDYNSALLSPSPS
FTLQSLSLVLDLISDRLIIAITGSLPQLVKLDLEDRPEKEPFPDNDLTYTGLQALGFC
QQLTSLSLVRTCYNRKISFKRINDMGIFLLSEACKGLESVRLGGFPKVSDAGFASLLH
SCRNLKKFEVRGAFLLSDLAFHDVTGSSCSLQEVRLSTCPLITSEAVKKLGLCGNLEV
LDLGSCKSISDSCLNSVSALRKLTSLNLAGADVTDSGMLALGKSDVPITQLSLRGCRR
VSDRGISYLLNNEGTISKTLSTLDLGHMPGISDRAIHTITHCCKALTELSIRSCFHVT
DSSIESLATWERQAEGGSKQLRKLNVHNCVSLTTGALRWLSKPSFAGLHWLGMGQTRF
AGRKETVTAMICGQRPWLTLCFDGCELGCSDGWEFHTPQRH"
CDS complement(join(12184. .12661,12760. .12912,12999. .13311,
13398. .13599,14054. .14428))
/note="hypothetical protein"
/codon_start=1
/evidence=not_experimental
/product="F20N2.3"
/protein_id="AAF79511.1"
/db_xref="GI:8778503"
/translation="MSDFDENFIEMTSYWAPPSSPSPRTILAMLEQTDNGLNPISEIF
PQESLPRDHTDQSGQRSGLRERLAARVGFNLPTLNTEENMSPLDAFFRSSNVPNSPVV
AISPGFSPSALLHTPNMVSDSSQGSLGLNSKTVTKIVFVTFLLQIIPPSSATNYGPLE
MVETSGEDNAAMMFNNDLPYQPYNVDLPSLEVFDDIATEESFYIPSYEPHVDPIGTP
LVTSFESELVDDAHTDIISIEDSESEDGNKDDDDEDFQYEDEDEDQYDQDQDVDEDEE
EEKDEDNVALDDPQPPPPKRRRYEVSNMIGATRTSKTQRIILQMESDEDNPNDGYRWR
KYGQKVVKGNPNPRSYFKCTNIECRVKKHVERGADNIKLVVTTYDGIHNHPSPPARRS
NSSSRNRSAGATIPQNQNDRTSRLGRAPPTPTPPTPPPSSYTPEEMRPFSSLATEIDL
TEVYMTGISMLPNIPVYENSGFMYQNDEPTMNAMPDGSDVYDGIMERLYFKFGVDM"
CDS complement(18009. .21509)
/note="hypothetical protein"
/codon_start=1
/evidence=not_experimental
/product="F20N2.4"
/protein_id="AAF79510.1"
/db_xref="GI:8778502"
/translation="MKQRWLLVLILCFFTTSLVMGIHGKHLINDDFNETALLLAFKQN
SVKSDPNNVLGNWKYESGRGSCSWRGVSCSDDGRIVGLDLRNSGLTGTLNLVNLTALP
NLQNLYLQGNYFSSGGDSSGSDCYLQVLDLSSNSISDYSMVDYVFSKCSNLVSVNISN
NKLVGKLGFAPSSLQSLTTVDLSYNILSDKIPESFISDFPASLKYLDLTHNNLSGDFS
DLSFGICGNLTFFSLSQNNLSGDKFPITLPNCKFLETLNISRNNLAGKIPNGEYWGSF
QNLKQLSLAHNRLSGEIPPELSLLCKTLVILDLSGNTFSGELPSQFTACVWLQNLNLG
NNYLSGDFLNTVVSKITGITYLYVAYNNISGSVPISLTNCSNLRVLDLSSNGFTGNVP
SGFCSLQSSPVLEKILIANNYLSGTVPMELGKCKSLKTIDLSFNELTGPIPKEIWMLP
NLSDLVMWANNLTGTIPEGVCVKGGNLETLILNNNLLTGSIPESISRCTNMIWISLSS
NRLTGKIPSGIGNLSKLAILQLGNNSLSGNVPRQLGNCKSLIWLDLNSNNLTGDLPGE
LASQAGLVMPGSVSGKQFAFVRNEGGTDCRGAGGLVEFEGIRAERLERLPMVHSCPAT
RIYSGMTMYTFSANGSMIYFDISYNAVSGFIPPGYGNMGYLQVLNLGHNRITGTIPDS
FGGLKAIGVLDLSHNNLQGYLPGSLGSLSFLSDLDVSNNNLTGPIPFGGQLTTFPVSR
YANNSGLCGVPLRPCGSAPRRPITSRIHAKKQTVATAVIAGIAFSFMCFVMLVMALYR
VRKVQKKEQKREKYIESLPTSGSCSWKLSSVPEPLSINVATFEKPLRKLTFAHLLEAT
NGFSAETMVGSGGFGEVYKAQLRDGSVVAIKKLIRITGQGDREFMAEMETIGKIKHRN
LVPLLGYCKVGEERLLVYEYMKWGSLETVLHEKSSKKGGIYLNWAARKKIAIGAARGL
AFLHHSCIPHIIHRDMKSSNVLLDEDFEARVSDFGMARLVSALDTHLSVSTLAGTPGY
VPPEYYQSFRCTAKGDVYSYGVILLELLSGKKPIDPGEFGEDNNLVGWAKQLYREKRG
AEILDPELVTDKSGDVELFHYLKIASQCLDDRPFKRPTMIQLMAMFKEMKADTEEDES
LDEFSLKETPLVEESRDKEP"
CDS complement(join(25636. .26376,26678. .26782,26909. .27363,
27581. .27775,28180. .28582,28672. .29127))
/note="hypothetical protein"
/codon_start=1
/evidence=not_experimental
/product="F20N2.5"
/protein_id="AAF79509.1"
/db_xref="GI:8778501"
/translation="MSSGGAGEYNEDRHLLRSTDGDEVGIGGGEGDLDVESQSPAIRS
GAGGVRDLFKHIDRRFSLSGRRLSFKRMENIRVDRERHNPSSSSAFSAAGEEDGGGIS
NLHSVDDRNDEYGFDEEVLGDSAPPEWALLLIGCLIGVAAGICVAGFNKGVHVIHEWA
WAGTPNEGAAWLRLQRLADTWHRILLIPVTGGVIVGMMHGLLEILDQIRQSNSSQRQG
LDFLAGIYPVIKAIQAAVTLGTGCSLGPEGPSVDIGKSCANGFALMMENNRERRIALT
AAGAASGIASGFNAAVAGCFFAIETVLRPLRAENSPPFTTAMIILASVISSTVSNALL
GTQSAFTVPSYDLKSAAELPLYLILGMLCGAVSVVFSRLVTWFTKSFDFIKDKFGLPA
IVCPALGGLGAGIIALKYPGILYWGFTNVEEILHTGKSASAPGIWLLAQLAAAKVVAT
ALCKGSGLVGGLYAPSLMIGAAVGAVFGGSAAEIINRAIPGNAAVAQPQAYALVGMAA
TLASMCSVPLTSVLLLFELTKDYRILLPLMGAVGLAIWVPSVANQGKESDSSEGRSTG
RGYSSLSPSERKTEGVWRHTDNADSLELTVIENPDHNSFLDEETILEDLKVMRVMSKN
YVKVSSGTTLREARNILKESHQNCIMVVDDDDFLAGILTHGDIRRYLSNNASTILDVS
SETKQTPSVSSHLVEPSLSLRLLQENTCPVSSVCTKKISYRGQERGLLTCYPDATVGV
AKELMEARGVKQLPVVKRGEVIHKGKRRKLLGLLHYDSIWTFLR"
tRNA 29448. .29529
/note="codon recognized: TCA; putative tRNA-Ser"
/product="tRNA-Ser"
/evidence=not_experimental
CDS complement(join(29954. .31341,31444. .31529,31594. .31784))
/note="hypothetical protein"
/codon_start=1
/evidence=not_experimental
/product="F20N2.6"
/protein_id="AAF79508.1"
/db_xref="GI:8778500"
/translation="MGQFYFRAHKSRVGSSSKCGFGSAILFSVVLKSREQRQRRLSRT
WDPGSVILPLYPLNSSSSSSLMGFCGFANPRLSEVRFSSLNLEFSGLLTSRFLFTSRK
FCNGSIGGDVTDNGTEEPLKITWESSEMDCEFDQEENGEKISVRKRFMESTKLSASRV
LDTLQQDCPGFNTKSALDELNVSISGLLVREVLVGILRTLSFDNKTRCAKLAYKFFVW
CGGQENFRHTANCYHLLMKIFAECGEYKAMCRLIDEMIKDGYPTTACTFNLLICTCGE
AGLARDVVEQFIKSKTFNYRPYKHSYNAILHSLLGVKQYKLIDWVYEQMLEDGFTPDV
LTYNIVMFANFRLGKTDRLYRLLDEMVKDGFSPDLYTYNILLHHLATGNKPLAALNLL
NHMREVGVEPGVIHFTTLIDGLSRAGKLEACKYFMDETVKVGCTPDVVCYTVMITGYI
SGGELEKAEEMFKEMTEKGQLPNVFTYNSMIRGFCMAGKFKEACALLKEMESRGCNPN
FVVYSTLVNNLKNAGKVLEAHEVVKDMVEKGHYVHLISKLKKYRRS"
CDS 31908. .32471
/note="hypothetical protein"
/codon_start=1
/evidence=not_experimental
/product="F20N2.7"
/protein_id="AAF79494.1"
/db_xref="GI:8778486"
/translation="MLAPGESVLIPAEEISLSAGDVISLSVHNLIASVSSYRPWWSEF

Query Match 17.0%; Score 294.2; DB 12; Length 109171;
Best Local Similarity 56.3%; Pred. No. 1.1e-52;
Matches 719; Conservative 0; Mismatches 483; Indels 75; Gaps 6;

Qy 222 ctactgctacaactatcaccaccaccaccatcactacctcaccagctattcaaatccgcc 281
Db 2304 CTCCTCCTTTGTGCATCGCCTCCTCCTCGGCCGCAACCTCCGCCTCACATCACCTCCGTC 2363

Qy 282 agctactcattagctgtgcggagttgatttcgcagtccgatttctcggccgcgaaaagac 341
Db 2364 GTCTTCTTTTCACCGCTGCGAATTTCGTCTCCCAGTCAAACTTCACCGCCGCTCAAAACT 2423

```
Qy    342 tccttactatattatcaactaactcatctccttttggtgattcaactgaacggttagtcc 401
          | ||  | ||  | ||   |||||| ||||||   || || || || || ||  | || |
Db   2424 TACTCTCAATCCTCTCCCTTAACTCTTCTCCTCACGGCGACTCCACCGAGCGACTTGTAC 2483

Qy    402 atcaatttactcgcgcactttcccttcgtctcaaccgcta---------- 441
          | |  || |||   ||  | ||| | ||  |||||||  |
Db   2484 ACCTCTTCACTAAAGCCTTGTCCGTACGAATCAACCGTCAGCAACAAGATCAGACGGCTG 2543

Qy    442 -------tatatcgtcaaccaccaatcatttcatgacacctgttgaaacaac-------- 486
               |   |||  || || ||  |  | | ||    |      ||
Db   2544 AAACGGTTGCCACGTGGACGACGAACGAAATGACGATGAGTAACTCCACGGTGTTCACGA 2603

Qy    487 --------tccaactgattcttcttcttcgtcatcattagctctaattcaatcatc---- 534
                | |||   |  ||||| ||    | ||  | |  |||||   | ||
Db   2604 GCAGTGTATGCAAAGAACAGTTCTTGTTTCGAACCAAGAACAACAATTCTGACTTCGAGT 2663

Qy    535 -------atatctatctctaaaccaagttacccctttcataaggtttactcaattaaccg 587
               |||| |  |||||||||| | || || || ||  ||||   ||| |||| |
Db   2664 CTTGTTACTATCTTTGGCTAAACCAACTAACGCCGTTTATTCGGTTCGGTCATTTAACGG 2723

Qy    588 ctaatcaagcgattttagaagcgattaacggtaatcatca---agcaatccacatcgttg 644
          | || ||||| ||  | || ||||   |    ||  || |   |||  | || ||  |||
Db   2724 CGAACCAAGCTATCCTCGACGCGACGGAGACAAACGATAACGGAGCTCTACATATACTTG 2783

Qy    645 atttcgacattaatcacggggttcaatggccaccgttaatgcaagcactagctgatcgtt 704
          |||| || ||    || ||  |||||||||| || || |||||| ||||| ||  | |
Db   2784 ATTTAGATATATCACAAGGACTTCAATGGCCTCCATTGATGCAAGCCCTAGCAGAGAGGT 2843

Qy    705 ac---------------cctgctcccactcttcgaatcaccggtactggaaatgaccttg 749
                          ||  |||| |||| || || |||||    ||    ||  |
Db   2844 CATCAAACCCTAGCAGTCCACCTCCATCTCTCCGCATAACCGGATGCGGTCGAGATGTAA 2903

Qy    750 ataccccttcgtagaacaggtgatcgtttagctaaatttgctcactcattagggttgagat 809
               |    |||| || || || ||| |    || ||| |||| |||||  |   ||
Db   2904 CCGGATTAAACCGAACTGGAGACCGGTTAACCCGGTTCGCTGACTCTTTAGGTCTCCAAT 2963

Qy    810 ttcaattccatcctctttatatagccaataataaccacgatcacgatgaagatccttcta 869
          | || || ||  | ||    || |   |  | | | ||       |      |
Db   2964 TCCAGTTTCACACGCTAGTGATCGTAGAAGAAGATCTCGCCGGACTTTTGCTACAGATCC 3023

Qy    870 ttatttcctccattgtactactccctgatgaaaccctagctatcaactgtgttttctacc 929
             | |   |  |   |    | |  |  || ||| | ||  |||| ||||||  || ||
Db   3024 GATTGTTAGCTCTCTCAGCCGTACAAGGAGAGACCATTGCCGTCAATTGTGTTCACTTCC 3083

Qy    930 tccaccgccttttaaaagac---cgcgaaaagttaaggatttttttgcatagggttaagt 986
          |||||    | || || |||     | || | | |    || |||     | | |||
Db   3084 TCCACAAAATATTTAACGACGATGGAGATATGATCGGTCACTTCTTGTCAGCGATCAAGA 3143

Qy    987 caatgaaccctaaaattgttacaatcgcggagaaggaagcaaatcataaccatcctcttt 1046
             | ||| ||| ||| |||||||| || ||||  |||||| ||||||    |||     |
Db   3144 GCTTAAACTCTAGAATCGTTACAATGGCAGAGAGAGAAGCTAATCATGGAGATCACTCGT 3203

Qy   1047 ttttacaaagattcatcgaggcgttggattattatacagctgtgtttgattcactggaag 1106
          | ||  | ||||||   |||||  ||||| |||| |  ||  | ||||||||  ||||||
Db   3204 TCTTGAATAGATTCTCTGAGGCAGTGGATCATTACATGGCGATCTTTGATTCGTTGGAAG 3263

Qy   1107 ctacattgccaccgggtagtcgagagaggatgacagttgaacaagtgtggtttgggagag 1166
          | || ||||| ||   ||| |||||||  | ||  | || |||  |||||| || |  |
Db   3264 CGACGTTGCCGCCAAATAGCCGAGAGAGACTAACCCTAGAGCAACGGTGGTTCGGTAAGG 3323

Qy   1167 agattgttgatatcgttgcgatggaaggagataaaaggaaagaaagacatgaaaggttta 1226
          |||||  | ||| | || |||  |||||       | || ||  |||||||   ||||||
Db   3324 AGATTTTGGATGTTGTGGCGGCGGAAGAGACGGAGAGAAAGCAAAGACATCGGAGGTTTG 3383

Qy   1227 gatcatgggaagttatgttgaggagttgtggatttagtaatgttgctttaagccctttttg 1286
               |||||||  ||| ||| ||| | ||| ||   ||| ||| || |  |    ||||
Db   3384 AGATTTGGGAAGAGATGATGAAGAGGTTTGGTTTCGTTAACGTTCCTATTGGAAGCTTTG 3443

Qy   1287 cattatcacaagctaagcttcttttgagacttcattatccttctgaaggctatcaactcg 1346
          | || || |||||||||||||| | |||||||||||||||| ||||| ||| | ||
Db   3444 CTTTGTCTCAAGCTAAGCTTCTTCTTAGACTTCATTATCCTTCAGAAGGTTATAATCTTC 3503
```

```
Qy   1347 gagtttcgagtaattctttcttcttaggttggcaaaatcaaccccttttctccatctcgt 1406
               |    |  |||||||| ||  | || ||||||||||  ||||| |||||| | ||||
Db   3504 AGTTCCTTAACAATTCTTTGTTTCTTGGCTGGCAAAATCGTCCCCTCTTCTCCGTTTCGT 3563

Qy   1407 cttggcgttgagaaaaa 1423
          | |||   |||   |||
Db   3564 CGTGGAAATGATCCAAA 3580
```

```
RESULT  7
CRU303349
LOCUS       CRU303349   27056 bp    DNA             PLN       10-APR-2001
DEFINITION  Capsella rubella ORF1, ORF2, ORF3, ORF4, ORF5 and ORF6 (partial).
ACCESSION   AJ303349
VERSION     AJ303349.1  GI:13620165
KEYWORDS    ORF1; ORF2; ORF3; ORF4; ORF5; ORF6.
SOURCE      Capsella rubella.
  ORGANISM  Capsella rubella
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Capsella.
REFERENCE   1  (bases 1 to 27056)
  AUTHORS   Rossberg,M., Theres,K., Acarkan,A., Herrero,R., Schmitt,T.,
            Schumacher,K., Schmitz,G. and Schmidt,R.
  TITLE     Comparative sequence analysis reveals extensive microcolinearity in
            the lateral suppressor regions of the tomato, arabidopsis, and
            capsella genomes
  JOURNAL   Plant Cell 13 (4), 979-988 (2001)
  MEDLINE   21178822
REFERENCE   2  (bases 1 to 27056)
  AUTHORS   Schmidt,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JAN-2001) Schmidt R.H., Max-Delbrueck-Laboratorium in
            der Max-Planck-Gesellschaft, Max-Delbrueck-Laboratorium in der
            Max-Planck-Gesellschaft, Carl-von-Linne-Weg 10, 50829 Koeln,
            GERMANY
FEATURES             Location/Qualifiers
     source          1. .27056
                     /organism="Capsella rubella"
                     /db_xref="taxon:81985"
                     /clone_lib="cosmid library"
     exon            1616. .2959
                     /note="ORF1"
                     /number=1
     CDS             1616. .2959
                     /note="ORF1"
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="CAC36387.1"
                     /db_xref="GI:13620166"
                     /translation="MLTSFKSSSSSSEDATTETPPPPFCQLASSSSSSAATSASHHLR
                     RLLFTAADFVSQSNFTAARNLLSILSLNSSPYGDSTQRLAHLFTKALSLRINRLQQEQ
                     DPTVATCTTNEMTMSTNSTVFTSSVCKEQFLFRTKNNNSDFESCYYLWLNQLTPFIRF
                     GHLTANQAILDATETNDNGALHILDLDISQGLQWPPLMQALAERSSSPNSPPPSLRIT
                     GCGRDVTGLNRTGDRLTRFANSLGLQFQFHKLVIVDEDLPGLLLQIRLLALSAVQGET
                     IAVNCVHFLYKFFNDDADLIGHFLTVIKSLNPRIVTMAEREANHGDHSFLIRFSEALD
                     HYVAIFDSLEATLPPNSRERLTLEQRWFGKEIMDVVAAEATERKQRHRRFEIWGEMMK
                     RFGFVNVPIGSFALSQAKLLLRLHYPSEGYNLQFLNDSLFLGWQNRLLFSVSSWK"
     mRNA            1616. .2959
                     /note="ORF1"
     mRNA            complement(join(7709. .9082,9175. .9265,9593. .9798,
                     9871. .10020))
                     /note="ORF2"
     CDS             complement(join(7709. .9082,9175. .9265,9593. .9798,
                     9871. .10020))
                     /note="ORF2"
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="CAC36388.1"
                     /db_xref="GI:13620167"
                     /translation="MEEVTRSVLAEEILKRLDLENLCSVACVSTTLRSAVVSGVLPSL
                     TSLDLSVFSPDEETLNHVLRGCIGLSSLTLNCLRLDANSVRGVLGPHLRELHLLRCSL
                     LSSTVLTSIGTLCPNLRVLTLEMSDLDSPAVFQSNLTQMLNGCPYLESLQLNIRGILV
```

```
                DATAFQSVRFSLPETLKALRLQPLLESEAILLMNRFKVTGTCLSQADYSALLSLSPSF
                TLQSLSLVLDLISDRLIIAITGSLPQLVQLDLEDRPEKEPFPDSDLTYTGLQALGYCQ
                QLTSLSLVRTCYNRKISFKRINDMGIFLLSEACKGLESVRLGGFQKVSDAGFASLLHS
                CRNLKKFEVRGAFLLSDLAFHDVTGSSCFLQEVRLSTCPLITSEAVKKLGLCGNLEVL
                DLGSCKSISDSCLNSVSVLRRLTSLNLAGADVTDSGMLALGKSDVPITQLSLRGCKRV
                SDRGISHLLINEGTISKTLSTLDLGHMPGMSDRAIHTITRYCKALTELSIRSCFYVTD
                SSIESLATRERQVEGGSKQLRKLNIHNCVSLTTGALRWLSKPSFAGLHWLGLGQTRIA
                GRKETVTATICGQRPWLTLCFDGCELGCYDGWEFHTPQRH"
exon            complement(7709. .9082)
                /note="ORF2"
                /number=4
intron          complement(9083. .9174)
                /note="ORF2"
                /number=3
exon            complement(9175. .9265)
                /note="ORF2"
                /number=3
intron          complement(9266. .9592)
                /note="ORF2"
                /number=2
exon            complement(9593. .9798)
                /note="ORF2"
                /number=2
intron          complement(9799. .9870)
                /note="ORF2"
                /number=1
exon            complement(9871. .10020)
                /note="ORF2"
                /number=1
exon            complement(10491. .10953)
                /note="ORF3"
                /number=5
mRNA            complement(join(10491. .10953,11109. .11264,11359. .11761,
                11859. .12000,12520. .12897))
                /note="ORF3"
CDS             complement(join(10491. .10953,11109. .11264,11359. .11761,
                11859. .12000,12520. .12897))
                /note="ORF3"
                /codon_start=1
                /product="hypothetical protein"
                /protein_id="CAC36389.1"
                /db_xref="GI:13620168"
                /translation="MSDFDGNFNESTSCWAPPFSPSPRAILAMIDQAEDDGLDLISEM
                FPHTNLPRDHSAQSSEQRSGLGQRLAARVGFSPPPLDTENISPSAASLRSSTTVPPPV
                SAIPPGFIPSTLLQSPNMVTDSSQMIPAPSSDTNYGPQGMVETSGEGNAPTLIFNNDH
                PYQPVNADLPLLDVFDDVPTEESVYLPTYETDHVEPIGDSLVPNSFEHEFVGDANVNF
                ISNEGTESVKDSESDDELDFDFEFEDEDEAEAEAEDNDEDKAEDEDEDNDEDGNEDED
                EDDEEDENDEEEADHDDEADFDDAEPSSPKRRKFEASSNMIGATRTNKAQRVILQMES
                DEDNPEDGFRWRKYGQKVVKGNPNPRSYFKCTNNDCNVKKHVERGADNFKILVTSYDG
                IHNHPPPPARCRINSGPRNRSGTTTTTQNQSYRTDRLGRFPAPSSVITPMEMMPLSSF
                TPPLDMSKVYMTGLNKLPNFSVFQNPGYMYRNDEPMMNVVPNGSGIYGGIRHRLFAEF
                GVNLM"
intron          complement(10954. .11108)
                /note="ORF3"
                /number=4
exon            complement(11109. .11264)
                /note="ORF3"
                /number=4
intron          complement(11265. .11358)
                /note="ORF3"
                /number=3
exon            complement(11359. .11761)
                /note="ORF3"
                /number=3
intron          complement(11762. .11858)
                /note="ORF3"
                /number=2
exon            complement(11859. .12000)
                /note="ORF3"
                /number=2
intron          complement(12001. .12519)
                /note="ORF3"
                /number=1
exon            complement(12520. .12897)
                /note="ORF3"
                /number=1
mRNA            complement(15045. .18545)
                /note="ORF4"
CDS             complement(15045. .18545)
                /note="ORF4"
                /codon_start=1
                /product="hypothetical protein"
                /protein_id="CAC36390.1"
                /db_xref="GI:13620169"
                /translation="MKQKWLFVLILCFFTALGIHGKRLINSDFDETALLMAFKQFSVK
                SDPNNVLGNWIYESGRGSCSWRGVSCSDDGRIVGLDLRNGGVTGTLNLANLTALPNLQ
                NLYLQGNYFSSSSGGDSSSGSYCYLQVLDLSSNLISDYSLVDYVFSKCSNLVSVNFSN
                NKLVGKLGFAPSSLKSLTTVDFSYNILSEKIPESFISEFPASLKYLDLTHNNFSGDFS
                DLSFGMCGNLSFFSLSQNNISGVKFPISLPNCRFLETLNISRNNLAGKIPGGEYWGSF
                QNLKQLSLAHNRFSGEIPPELSLLCKTLETLDLSGNALSGELPSQFTACVWLQNLNIG
                NNYLSGDFLSTVVSKITRITYLYVAFNNISGSVPISLTNCTNLRVLDLSSNGFTGNVP
                SGLCSQQSSPVLEKLLIANNYLSGTVPVELGKCKSLKTIDLSFNELTGPIPKDVWMLP
                NLSDLVMWANNLTGSIPEGVCVKGGKLETIILNNNLLTGSIPQSISRCTNMIWISLSS
                NRLTGKIPTGIGNLSKLAILQLGNNSLSGNVPRQLGNCKSLIWLDLNSNNLTGDLPGE
                LASQAGLVMPGSVSGKQFAFVRNEGGTDCRGAGGLVEFEGIRAERLERFPMVHSCPAT
                RIYSGMTMYTFSANGSMIYFDISYNAVSGLIPPGYGNMGYLQVLNLGHNRITGNIPDS
                LGGLKAIGVLDLSHNDLQGYLPGSLGSLSFLSDLDVSNNNLTGPIPFGGQLTTFPVSR
                YANNSGLCGVPLRPCGSAPRRPITSSVHAKKQTLATAVIAGIAFSFMCLVMLFMALYR
                VRKVQKKELKREKYIESLPTSGSCSWKLSSVPEPLSINVATFEKPLRKLTFAHLLEAT
                NGFSAETMVGSGGFGEVYKAQLRDGSVVAIKKLIRITGQGDREFMAEMETIGKIKHRN
                LVPLLGYCKVGEERLLVYEYMKWGSLETVLHEKSSKKGGIFLNWTARKKIAIGAARGL
                AFLHHSCIPHIIHRDMKSSNVLLDEDFEARVSDFGMARLVSALDTHLSVSTLAGTPGY
                VPPEYYQSFRCTAKGDVYSYGVILLELLSGKKPIDPGEFGEDNNLVGWAKQLYREKSG
                TEILDPELVTEKSGDAELFHYLKIASQCLDDRPFKRPTMIQVMAMFKELKADTEEDES
                LDEFSLKETPLVEESRDKEP"
exon            complement(22144. .22216)
                /note="ORF5"
                /number=8
mRNA            complement(join(22144. .22216,22307. .22539,22641. .23066,
                23369. .23473,23598. .24052,24269. .24463,24977. .25379,
                25500. .25952))
                /note="ORF5"
CDS             complement(join(22144. .22216,22307. .22539,22641. .23066,
                23369. .23473,23598. .24052,24269. .24463,24977. .25379,
                25500. .25952))
                /note="ORF5"
                /codon_start=1
                /product="hypothetical protein"
                /protein_id="CAC36391.1"
                /db_xref="GI:13620170"
                /translation="MSSGGAGEYNEDRHLLRSTGDEDSIGREDGDIDVESQSPAVRSG
                AGGVRDLFKHLDRRFSLSGRRLSFKRMENNRVDRERHNPSSSSSAFSAAEEDGGGISN
                LHNVDDRIDEYGFDEEVLGDSAPPEWALLLIGCLIGVAAGICVAGFNKGVHVIHEWAW
                AGTPNEGAAWLRLQRLADTWHRILLIPVTGGVIVGMMHGLLEILDQIRQSTSSQRQGL
                DFLAGIYPVIKAIQAAVTLGTGCSLGPEGPSVDIGKSCANGFALMMENNRERRIALTA
                AGAASGIASGFNAAVAGCFFAIETVLRPLRAENSPPFTTAMIILASVISSTVSNALLG
                TQSAFTVPSYDLKSAAELPLYLILGMLCGAVSVVFSRLVTWFTKSFDFIKEKFGLPAI
                VCPALGGLGAGIIALKYPGILYWGFTNVEEILHTGKSASAPGIWLLAQLAAAKVVATA
                LCKGSGLVGGLYAPSLMIGAAVGAVFGGSAAEIINRAIPGNAAVAQPQAYALVGMAAT
                LASMCSVPLTSVLLLFELTKDYRILLPLMGAVGLAIWVPSVANQGKESESSEGRNTGR
                GYSSVSPLERKTEGVWRHTDNVDSLELTVIESPDHKSFLDEETILEDLKVLRVMSKNY
                VKVSPGMTLREARNILKDSHQNCLMVVDEDDFLAGILTHGDIRRYLSNNVSTILDENT
                CQVSSVCTKNIIYRGQERGLLTCYPDATVGVAKELMEARGVKQLPVVKRGEVIHKGKR
                RKLLGLLHYDSIWTFLRDEMSRRRSINDRRKDKEVGANGH"
intron          complement(22217. .22306)
                /note="ORF5"
                /number=7
exon            complement(22307. .22539)
```

Query Match 15.9%; Score 275.2; DB 14; Length 27056;
Best Local Similarity 58.6%; Pred. No. 1.2e-48;
Matches 544; Conservative 0; Mismatches 363; Indels 21; Gaps 3;

```
Qy    536 tatctatctctaaaccaagttacccctttcataaggtttactcaattaaccgctaatcaa 595
          ||||| |  || || ||| | || || || ||  ||||   ||| ||||| || || ||
Db   2057 TATCTTTGGCTGAATCAACTAACGCCGTTTATTCGGTTCGGTCATTTAACGGCGAACCAG 2116
```

```
Qy    596 gcgattttagaagcgattaacggtaatcatca---agcaatccacatcgttgatttcgac 652
          |||||  | || ||||   |    ||  || |   |||  | || ||  ||||||| ||
Db   2117 GCGATCCTCGACGCGACCGAGACAAACGATAATGGAGCTCTACATATACTTGATTTAGAT 2176

Qy    653 attaatcacggggttcaatggccaccgttaatgcaagcactagctgatcgtt-------- 704
          ||    || ||  |||| |||||  || || |||||||| |||||  ||  | |
Db   2177 ATATCACAAGGACTTCAGTGGCCTCCATTGATGCAAGCCCTAGCAGAGAGGTCATCAAGC 2236

Qy    705 -------accctgctcccactcttcgaatcaccggtactggaaatgaccttgataccctt 757
                 ||  | ||  |||| || || |||||    ||    ||  |        |
Db   2237 CCTAACAGTCCACCACCATCTCTCCGCATAACCGGATGCGGTCGAGATGTAACCGGATTA 2296

Qy    758 cgtagaacaggtgatcgtttagctaaatttgctcactcattagggttgagatttcaattc 817
            |||| || ||| | ||| |    |||||| |||| |||||  |    || || ||
Db   2297 AACCGAACCGGAGATAGGTTAACCCGGTTTGCTAACTCTTTAGGTCTCCAGTTCCAGTTT 2356

Qy    818 catcctctttatatagccaataataaccacgatcacgatgaagatccttctattatttcc 877
          ||    ||    || |   |  | | | | |  |           |          | |
Db   2357 CACAAGCTGGTGATCGTAGACGAAGATCTCCCTGGGCTCTTGCTACAGATCCGATTGTTA 2416

Qy    878 tccattgtactactccctgatgaaaccctagctatcaactgtgttttctacctccaccgc 937
           |  |         | |  |  || ||| | ||  |||||||||||  || |||| ||
Db   2417 GCTCTCTCCGCCGTACAAGGAGAGACCATTGCCGTCAACTGTGTTCACTTCCTCTACAAG 2476

Qy    938 cttttaaaagaccgcgaaaagttaa---ggatttttttgcatagggttaagtcaatgaac 994
           | || || |||   | | | || |   |    || ||      | | |||    | |||
Db   2477 TTCTTTAACGACGATGCAGATTTGATCGGTCACTTCTTAACAGTGATCAAGAGCTTAAAC 2536

Qy    995 cctaaaattgttacaatcgcggagaaggaagcaaatcataaccatcctctttttttacaa 1054
          |||| ||| || ||||| || ||||  ||||| ||||||    |||     || ||
Db   2537 CCTAGAATCGTGACAATGGCTGAGAGAGAAGCTAATCATGGAGATCACTCGTTCTTGATT 2596

Qy   1055 agattcatcgaggcgttggattattatacagctgtgtttgattcactggaagctacattg 1114
          ||||||   ||||| |||||| |||    ||  | ||||||||  |||||||| || |||
Db   2597 AGATTCTCAGAGGCTTTGGATCATTACGTGGCGATATTTGATTCCTTGGAAGCGACTTTG 2656

Qy   1115 ccaccgggtagtcgagagaggatgacagttgaacaagtgtggtttgggagagagattgtt 1174
          || ||   ||| |||||||  | ||  | || |||  |||||| || |  |||||| |
Db   2657 CCGCCAAATAGCCGAGAGAGACTAACCCTAGAGCAACGGTGGTTCGGTAAGGAGATTATG 2716

Qy   1175 gatatcgttgcgatggaaggagataaaaggaaagaaagaacatgaaaggtttagatcatgg 1234
          ||| |||| |||  |||||      | |||||  |||||||   ||||||     ||||
Db   2717 GATGTCGTGGCGGCGGAAGCGACGGAGAGGAAGCAAAGACATCGGAGGTTTGAGATATGG 2776

Qy   1235 gaagttatgttgaggagttgtggatttagtaatgttgctttaagccctttgcattatca 1294
          | ||  ||| ||| ||| | ||| ||   ||| ||| || | |   ||||| || ||
Db   2777 GGAGAGATGATGAAGAGATTTGGTTTCGTTAACGTTCCTATTGGAAGCTTTGCTTTGTCT 2836

Qy   1295 caagctaagcttcttttgagacttcattatccttctgaaggctatcaactcggagtttcg 1354
          ||||||||||||||| | || || ||||||||||| ||||| ||| | ||    |
Db   2837 CAAGCTAAGCTTCTTCTTAGGCTCCATTATCCTTCAGAAGGTTATAATCTTCAGTTCCTT 2896

Qy   1355 agtaattctttcttcttaggttggcaaaatcaacccctttttctccatctcgtcttggcgt 1414
          |   ||||||| ||  | ||||||||||||||  |  || |||||| | || || |||
Db   2897 AACGATTCTTTGTTTCTTGGTTGGCAAAATCGTCTACTCTTCTCCGTTTCATCGTGGAAA 2956

Qy   1415 tgagaaaaactatcaaatagccaacttc 1442
          ||| | |||  |  ||  |   || | |
Db   2957 TGATACAAAGGAGGAACAAAAAAAATAC 2984


RESULT   8
A84084
LOCUS       A84084        687 bp    DNA             PAT       21-JAN-2000
DEFINITION  Sequence 13 from Patent WO9846759.
ACCESSION   A84084
VERSION     A84084.1  GI:6733223
KEYWORDS    .
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
            Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
            Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 687)
  AUTHORS   Theres,N.
  TITLE     PLANTS WITH CONTROLLED SIDE-SHOOT FORMATION AND/OR CONTROLLED
            ABSCISSION AREA FORMATION
  JOURNAL   Patent: WO 9846759-A 13 22-OCT-1998;
            THERES NIKOLAUS (DE)
FEATURES             Location/Qualifiers
     source          1. .687
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
BASE COUNT      184 a    146 c    168 g    189 t
ORIGIN


  Query Match           12.7%;  Score 218.8;  DB 9;  Length 687;
  Best Local Similarity   59.2%;  Pred. No. 9.8e-37;
  Matches  392;  Conservative    0;  Mismatches  267;  Indels    3;  Gaps    1;

Qy    707 cctgctcccactcttcgaatcaccggtactggaaatgaccttgataccttcgtagaaca 766
          ||  ||||  |||| || || |||||    ||    ||  |        |    ||||
Db     25 CCACCTCCATCTCTCCGCATAACCGGATGCGGTCGAGATGTAACCGGATTAAACCGAACT 84

Qy    767 ggtgatcgtttagctaaatttgctcactcattagggttgagatttcaattccatcctctt 826
          || || || ||| |    || ||| |||| |||||  |   ||| || || ||  | ||
Db     85 GGAGACCGGTTAACCCGGTTCGCTGACTCTTTAGGTCTCCAATTCCAGTTTCACACGCTA 144

Qy    827 tatatagccaataataaccacgatcacgatgaagatccttctattatttcctccattgta 886
            || |   |  |  | | ||        |       |          | |   |  |   |
Db    145 GTGATCGTAGAAGAAGATCTCGCCGGACTTTGCTACAGATCCGATTGTTAGCTCTCTCA 204

Qy    887 ctactccctgatgaaaccctagctatcaactgtgttttctacctccaccgccttttaaaa 946
            | |  |  || ||| | ||  |||| ||||||  || |||||||    | || ||
Db    205 GCCGTACAAGGAGAGACCATTGCCGTCAATTGTGTTCACTTCCTCCACAAAATATTTAAC 264

Qy    947 gac---cgcgaaaagttaaggatttttttgcatagggttaagtcaatgaaccctaaaatt 1003
          |||    | || | | |  |    || |||     | | |||    | ||| ||| |||
Db    265 GACGATGGAGATATGATCGGTCACTTCTTGTCAGCGATCAAGAGCTTAAACTCTAGAATC 324

Qy   1004 gttacaatcgcggagaaggaagcaaatcataaccatcctctttttttacaaagattcatc 1063
          |||||||| || ||||  |||||  |||||     |||     || ||  | ||||||
Db    325 GTTACAATGGCAGAGAGAGAAGCTAATCATGGAGATCACTCGTTCTTGAATAGATTCTCT 384

Qy   1064 gaggcgttggattattatacagctgtgtttgattcactggaagctacattgccaccgggt 1123
          |||||  |||||  |||| |  ||  | ||||||||  |||||||  || ||||| ||   |
Db    385 GAGGCAGTGGATCATTACATGGCGATCTTTGATTCGTTGGAAGCGACGTTGCCGCCAAAT 444

Qy   1124 agtcgagagaggatgacagttgaacaagtgtggtttgggagagagattgttgatatcgtt 1183
          || ||||||||  | ||  | || |||  ||||||  || |  ||||||  | |||  | ||
Db    445 AGCCGAGAGAGACTAACCCTAGAGCAACGGTGGTTCGGTAAGGAGATTTTGGATGTTGTG 504

Qy   1184 gcgatggaaggagataaaaggaaagaaagaacatgaaaggtttagatcatgggaagttatg 1243
          |||  |||||       | || ||  ||||||||   ||||||      |||||||  |||
Db    505 GCGGCGGAAGAGACGGAGAGAAAGCAAAGACATCGGAGGTTTGAGATTTGGGAAGAGATG 564

Qy   1244 ttgaggagttgtggatttagtaatgttgctttaagccctttgcattatcacaagctaag 1303
            ||| ||| | ||| ||   ||| ||| || |   |    ||||| || || |||||||||
Db    565 ATGAAGAGGTTTGGTTTCGTTAACGTTCCTATTGGAAGCTTTGCTTTGTCTCAAGCTAAG 624

Qy   1304 cttcttttgagacttcattatccttctgaaggctatcaactcggagtttcgagtaattct 1363
          |||||| | |||||||||||||||| ||||| ||| | ||      |    | ||||||
Db    625 CTTCTTCTTAGACTTCATTATCCTTCAGAAGGTTATAATCTTCAGTTCCTTAACAATTCT 684

Qy   1364 tt 1365
          ||
Db    685 TT 686


RESULT   9
ATRGA1
LOCUS       ATRGA1       2210 bp    mRNA            PLN       19-AUG-1997
DEFINITION  A.thaliana mRNA for RGA1 gene.
ACCESSION   Y11336
```

```
VERSION     Y11336.1  GI:2339977
KEYWORDS    RGA1 gene.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
            Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
            Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 2210)
  AUTHORS   Truong,H.N., Caboche,M. and Daniel-Vedele,F.
  TITLE     Sequence and characterization of two Arabidopsis thaliana cDNAs
            isolated by functional complementation of a yeast gln3 gdh1 mutant
  JOURNAL   FEBS Lett. 410 (2-3), 213-218 (1997)
  MEDLINE   97379310
REFERENCE   2  (bases 1 to 2210)
  AUTHORS   Truong,H.N.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-FEB-1997) H.N. Truong, INRA-Versailles, Laboratoire
            de Biologie Cellulaire, Route de Saint-Cyr, 78026 Versailles Cedex,
            FRANCE
FEATURES             Location/Qualifiers
     source          1. .2210
                     /organism="Arabidopsis thaliana"
                     /sub_species="(L). Heynh"
                     /db_xref="taxon:3702"
                     /dev_stage="two-leaf stage seedlings"
                     /note="Landsberg erecta ecotype"
     gene            132. .1895
                     /gene="RGA1"
     CDS             132. .1895
                     /gene="RGA1"
                     /codon_start=1
                     /product="RGA1 protein"
                     /protein_id="CAA72177.1"
                     /db_xref="GI:2339978"
                     /db_xref="SPTREMBL:O23642"
                     /translation="MKRDHHQFQGRLSNHGTSSSSSSISKDKMMMVKKEEDGGGNMDD
                     ELLAVLGYKVRSSEMAEVALKLEQLETMMSNVQEDGLSHLATDTVHYNPSELYSWLDN
                     MLSELNPPPLPASSNGLDPVLPSPEICGFPASDYDLKVIPGNAIYQFPAIDSSSSSNN
                     QNKRLKSCSSPDSMVTSTSTGTQIGGVIGTTVTTTTTTTTAAAESTRSVILVDSQENG
                     VRLVHALMACAEAIQQNNLTLAEALVKQIGCLAVSQAGAMRKVATYFAEALARRIYRL
                     SPPQNQIDHCLSDTLQMHFYETCPYLKFAHFTANQAILEAFEGKKRVHVIDFSMNQGL
                     QWPALMQALALREGGPPTFRLTGIGPPAPDNSDHLHEVGCKLAQLAEAIHVEFEYRGF
                     VANSLADLDASMLELRPSDTEAVAVNSVFELHKLLGRPGGIEKVLGVVKQIKPVIFTV
                     VEQESNHNGPVFLDRFTESLHYYSTLFDSLEGVPNSQDKVMSEVYLGKQICNLVACEG
                     PDRVERHETLSQWGNRFGSSGLAPAHLGSNAFKQASMLLSVFNSGQGYRVEESNGCLM
                     LGWHTRPLITTSAWKLSTAAH"
BASE COUNT      583 a    454 c    518 g    655 t
ORIGIN
```

```
  Query Match             8.2%;  Score 141.6;  DB 13;  Length 2210;
  Best Local Similarity   52.8%;  Pred. No. 3.3e-20;
  Matches  453;  Conservative    0;  Mismatches  384;  Indels   21;  Gaps    6;

Qy     560 cctttcataaggtttactcaattaaccgctaatcaagcgattttagaagcgattaacggt 619
           ||||   | |  ||  |||| || || || || |||||||||| | |||||  || | |||
Db    1029 CCTTATCTTAAATTCGCTCACTTCACGGCGAACCAAGCGATTCTCGAAGCTTTTGAAGGT 1088

Qy     620 aatcatcaagcaatccacatcgttgatttcgacattaatcacggggttcaatggccaccg 679
           ||  |   ||   | ||| || ||||||||   || || || ||  |||||||||| ||
Db    1089 AAGAAGAGAG---TACACGTCATTGATTTCTCGATGAACCAAGGTCTTCAATGGCCTGCG 1145

Qy     680 ttaatgcaagcactagctgatcgtta---ccctgctcccactcttcgaatcaccggtact 736
            | |||||||| || ||   |||  |      | |||| ||| | ||  | ||||| | |
Db    1146 CTTATGCAAGCTCTTGCGCTTCGAGAAGGAGGTCCTCCAACTTTCCGGTTAACCGGAATT 1205

Qy     737 gg---aaatgaccttgataccttcgtagaacaggtgatcgtttagctaaatttgctcac 793
           ||   |   |  | ||||    |  |         |||   |      |||||| |||||
Db    1206 GGTCCACCGGCGCCGGATAATTCTGATCATCTTCATGAAGTTGGTTGTAAATTAGCTCAG 1265

Qy     794 tcattagggttgagatttcaattccatcctctttatatagccaataataaccacgatcac 853
             | |  ||       |              |  |  |  |  ||||  |     |
Db    1266 CTTGCGGAGGCGATTCACGTAGAATTCGAATACCGTGGATTCGTTGCTAACAGCTTAGCC 1325

Qy     854 gatgaagatccttctattatttcctccattgtactactccctgatgaaaccctagctatc 913
           |||   ||| |||| ||  ||       ||  ||    |  |  ||| |  | ||  |
Db    1326 GATCTCGATGCTTCGATGCTT---GAGCTTAGACCGAGCGATACGGAAGCTGTTGCGGTG 1382

Qy     914 aactgtgttttctacctccaccgccttttaaaagaccgcgaaaagttaaggatttttttg 973
           |||| ||||||  | || ||     || |||       | ||    | ||  ||   || |
Db    1383 AACTCTGTTTTTGAGCTACATAAGCTCTTAGGTCGTCCCGGTGGGATAGAGAAAGTTCTC 1442

Qy     974 catagggttaagtcaatgaaccctaaaattgttacaatcgcggagaaggaagcaaatcat 1033
              || ||    || || ||    ||| | ||  | |  ||| | ||| | || |||
Db    1443 GGCGTTGTGAAACAGATTAAACCGGTGATTTTCACGGTGGTTGAGCAAGAATCGAACCAT 1502

Qy    1034 aaccatcctctttttttacaaagattcatcgaggcgttggattattatacagctgtgttt 1093
           |||   ||  |||| ||| |  | || |  ||  |||| |||||||| |  || |||||
Db    1503 AACGGACCGGTTTTCTTAGACCGGTTTACTGAATCGTTACATTATTATTCGACTCTGTTT 1562

Qy    1094 gattcactggaagctacattgccaccgggtagtcgagagaggatgacagttgaacaagtg 1153
           |||||  ||||||          |||  ||||| ||| |   | |    ||  ||||
Db    1563 GATTCGTTGGAAG------GAGTTCCGAATAGTCAAGACAAAGTCATGTCTG---AAGTT 1613

Qy    1154 tggtttgggagagagattgttgatatcgttgcgatggaaggagataaaaggaaagaaaga 1213
           |  || |||| | |||||  | || | || ||    |||||   | | ||    || |||
Db    1614 TACTTAGGGAAACAGATTTGTAATCTGGTGGCTTGTGAAGGTCCTGACAGAGTCGAGAGA 1673

Qy    1214 catgaaaggtttagatcatgggaagttatgttgaggagttgtggatttagtaatgttgct 1273
           || |||| ||| ||   ||||| |     |||  |    |  || ||       |   |
Db    1674 CACGAAACGTTGAGTCAATGGGGAAACCGGTTTGGTTCGTCCGGTTTAGCGCCGGCACAT 1733

Qy    1274 ttaagccctttttgcattatcacaagctaagcttcttttgagacttcattatccttctgaa 1333
            |  |  ||   || ||     ||||| |   | ||||||    |   | ||  |   ||
Db    1734 CTTGGGTCTAACGCGTTTAAGCAAGCGAGTATGCTTTTGTCTGTGTTTAATAGTGGCCAA 1793

Qy    1334 ggctatcaactcggagtttcgagtaattcttttcttcttaggttggcaaaatcaacccctt 1393
           || ||||   | |  |     | |   | |||  | || ||||||||  | ||  || ||
Db    1794 GGTTATCGTGTGGAGGAGAGTAATGGATGTTTGATGTTGGGTTGGCACACTCGCCCACTC 1853

Qy    1394 ttctccatctcgtcttgg 1411
            |  ||| |||  |||||
Db    1854 ATTACCACCTCCGCTTGG 1871
```

```
RESULT  10
ATY15194
LOCUS       ATY15194     2201 bp    mRNA.          PLN       01-NOV-1997
DEFINITION  Arabidopsis thaliana mRNA for GRS protein.
ACCESSION   Y15194
VERSION     Y15194.1  GI:2569939
KEYWORDS    GRS protein.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
            Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
            Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 2201)
  AUTHORS   Peng,J., Carol,P., Richards,D.E., King,K.E., Cowling,R.J.,
            Murphy,G.P. and Harberd,N.P.
  TITLE     The Arabidopsis GAI gene defines a signalling pathway that
            negatively regulates gibberellin responses
  JOURNAL   Genes Dev. In press
REFERENCE   2  (bases 1 to 2201)
  AUTHORS   Harberd,N.P.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-OCT-1997) N.P. Harberd, John Innes Centre, Molecular
            Genetics, Colney Lane, Norwich, NR4 7UJ, UK
FEATURES             Location/Qualifiers
     source          1. .2201
                     /organism="Arabidopsis thaliana"
                     /strain="Columbia"
                     /db_xref="taxon:3702"
                     /chromosome="2"
     CDS             144. .1907
                     /codon_start=1
```

```
                     /product="GRS protein"
                     /protein_id="CAA75493.1"
                     /db_xref="GI:2569940"
                     /db_xref="SPTREMBL:O23725"
                     /translation="MKRDHHQFQGRLSNHGTSSSSSSISKDKMMMVKKEEDGGGNMDD
                     ELLAVLGYKVRSSEMAEVALKLEQLETMMSNVQEDGLSHLAADTVHYNPSELYSWLDN
                     MLSELNPPPLPASSNGLDPVLPSPEICGFPASDYDLEVIPGNAIYQFPAIDSSSSSNN
                     QNKRLKSCSSPDSMVTSTSTGTQIGGVIGTTVTTTTTTTTAAAESTRSVILVDSQENG
                     VRLVHALMACAEAIQQNNLTLAEALVKQIGCLAVSQAGAMRKVATYFAEALARRIYRL
                     SPPQNQIDHCLSDTLQMHFYETCPYLKFAHFTANQAILEAFEGKKRVHVIDFSMNQGL
                     QWPALMQALALREGGPPTFRLTGIGPPAPDNSDHLHEVGCKLAQLAEAVHVEFEYRGF
                     VANSLADLDASMLELRPSDTEAVAVNSVFELHKLLGRPGGIEKVLGVVKQIKPVIFTV
                     VEQESNHNGPVFLDRFTESLHYYSTLFDSLEGVPNSQDKVMSEVYLGKQICNLVACEG
                     PDRVERHETLSQWGNRFGSSGLAPAHLGSNAFKQASMLLSVFNSGQGYRVEESNGCLM
                     LGWHTRPLITTSAWKLSTAAY"
BASE COUNT      576 a    450 c    525 g    650 t
ORIGIN


  Query Match             8.1%;  Score 140;  DB 14;  Length 2201;
  Best Local Similarity   52.7%;  Pred. No. 7.2e-20;
  Matches  452;  Conservative    0;  Mismatches  385;  Indels   21;  Gaps    6;

Qy     560 cctttcataaggtttactcaattaaccgctaatcaagcgattttagaagcgattaacggt 619
           ||||   | |  ||  |||| || || || || ||||||||| | |||||  || | |||
Db    1041 CCTTATCTTAAATTCGCTCACTTCACGGCGAACCAAGCGATTCTCGAAGCTTTTGAAGGT 1100

Qy     620 aatcatcaagcaatccacatcgttgatttcgacattaatcacggggttcaatggccaccg 679
           ||  |   ||   | ||| || |||||||||   || || || ||  ||||||||||  ||
Db    1101 AAGAAGAGAG---TACACGTCATTGATTTCTCGATGAACCAAGGTCTTCAATGGCCTGCG 1157

Qy     680 ttaatgcaagcactagctgatcgtta---ccctgctcccactcttcgaatcaccggtact 736
            | |||||||| || ||   |||  |      | |||| ||| | ||  | ||||| | |
Db    1158 CTTATGCAAGCTCTTGCGCTTCGAGAAGGAGGTCCTCCAACTTTCCGGTTAACCGGAATT 1217

Qy     737 gg---aaatgaccttgatacccttcgtagaacaggtgatcgtttagctaaatttgctcac 793
           ||   |   |  |  ||||    |  |        |||   |     ||||| |||||
Db    1218 GGTCCACCGGCGCCGGATAATTCTGATCATCTTCATGAAGTTGGTTGTAAATTAGCTCAG 1277

Qy     794 tcattagggttgagatttcaattccatcctctttatatagccaataataaccacgatcac 853
           | |  |       |              |  |  |  |  ||||  |     |
Db    1278 CTTGCGGAGGCGGTTCACGTAGAATTCGAATACCGTGGATTCGTTGCTAACAGCTTAGCC 1337

Qy     854 gatgaagatccttctattatttcctccattgtactactccctgatgaaaccctagctatc 913
           |||   ||| |||| ||  ||       ||  ||    |  |   ||| |  | ||  |
Db    1338 GATCTCGATGCTTCGATGCTTGAGC---TTAGACCGAGCGATACGGAAGCTGTTGCGGTG 1394

Qy     914 aactgtgttttctacctccaccgccttttaaaagaccgcgaaaagttaaggatttttttg 973
           |||| ||||||  | || ||    || |||      | ||    | ||  ||   || |
Db    1395 AACTCTGTTTTTGAGCTACATAAGCTCTTAGGTCGTCCCGGTGGGATAGAGAAAGTTCTC 1454

Qy     974 catagggttaagtcaatgaaccctaaaattgttacaatcgcggagaaggaagcaaatcat 1033
           ||  ||    || || ||    ||| | ||  | |  ||| | ||| | || |||
Db    1455 GGCGTTGTGAAACAGATTAAACCGGTGATTTTCACGGTGGTTGAGCAAGAATCGAACCAT 1514

Qy    1034 aaccatcctctttttttacaaagattcatcgaggcgttggattattatacagctgtgttt 1093
           |||   ||  |||| ||| |  | || |  ||  |||| |||||||| |  || |||||
Db    1515 AACGGACCGGTTTTCTTAGACCGGTTTACTGAATCGTTACATTATTATTCGACTCTGTTT 1574

Qy    1094 gattcactggaagctacattgccaccgggtagtcgagagaggatgacagttgaacaagtg 1153
           |||||  ||||||           |||  ||||| ||| |    | |    ||   ||||
Db    1575 GATTCGTTGGAAG------GAGTTCCGAATAGTCAAGACAAAGTCATGTCTG---AAGTT 1625

Qy    1154 tggtttgggagagagattgttgatatcgttgcgatggaaggagataaaaggaaagaaaga 1213
           |  || |||| | |||||  | || | || ||    |||||   | | ||    || |||
Db    1626 TACTTAGGGAAACAGATTTGTAATCTGGTGGCTTGTGAAGGTCCTGACAGAGTCGAGAGA 1685

Qy    1214 catgaaaggtttagatcatgggaagttatgttgaggagttgtggatttagtaatgttgct 1273
           || |||| ||| ||   ||||| |     |||  |    |  || ||       |    |
Db    1686 CACGAAACGTTGAGTCAATGGGGAAACCGGTTTGGTTCGTCCGGTTTAGCGCCGGCACAT 1745

Qy    1274 ttaagccctttttgcattatcacaagctaagcttcttttgagacttcattatccttctgaa 1333
           |  |  ||   || ||    ||||| |   | ||||||    |   | ||  |    ||

Db    1746 CTTGGGTCTAACGCGTTTAAGCAAGCGAGTATGCTTTTGTCTGTGTTTAATAGTGGCCAA 1805

Qy    1334 ggctatcaactcggagtttcgagtaattcttttcttcttaggttggcaaaatcaacccctt 1393
           || ||||   | | |     | |   | |||  | || ||||||||| | ||  || ||
Db    1806 GGTTATCGTGTGGAGGAGAGTAATGGATGTTTGATGTTGGGTTGGCACACTCGTCCACTC 1865

Qy    1394 ttctccatctcgtcttgg 1411
            |  ||| |||  |||||
Db    1866 ATTACCACCTCCGCTTGG 1883


RESULT  11
AC005560/c
LOCUS       AC005560    95137 bp    DNA             PLN       05-APR-2000
DEFINITION  Arabidopsis thaliana chromosome II section 3 of 255 of the complete
            sequence. Sequence from clones F10A8, F2I9.
ACCESSION   AC005560 AE002093
VERSION     AC005560.2  GI:6598472
KEYWORDS    HTG.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
            Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
            Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 95137)
  AUTHORS   Lin,X., Kaul,S., Rounsley,S.D., Shea,T.P., Benito,M.-I., Town,C.D.,
            Fujii,C.Y., Mason,T.M., Bowman,C.L., Barnstead,M.E.,
            Feldblyum,T.V., Buell,C.R., Ketchum,K.A., Lee,J.J., Ronning,C.M.,
            Koo,H., Moffat,K.S., Cronin,L.A., Shen,M., VanAken,S.E., Umayam,L.,
            Tallon,L.J., Gill,J.E., Adams,M.D., Carrera,A.J., Creasy,T.H.,
            Goodman,H.M., Somerville,C.R., Copenhaver,G.P., Preuss,D.,
            Nierman,W.C., White,O., Eisen,J.A., Salzberg,S.L., Fraser,C.M. and
            Venter,J.C.
  TITLE     Sequence and analysis of chromosome 2 of the plant Arabidopsis
            thaliana
  JOURNAL   Nature 402 (6763), 761-768 (1999)
  MEDLINE   20083487
   PUBMED   10617197
REFERENCE   2  (bases 1 to 95137)
  AUTHORS   Lin,X.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-MAR-2000) The Institute for Genomic Research, 9712
            Medical Center Dr., Rockville, MD 20850, USA
COMMENT     On Dec 17, 1999 this sequence version replaced gi:3785968.
            The sequence and annotation of chromosome 2 were merged from those
            of the individual clones on this chromosome after removing
            overlaps. For detailed information, please see the TIGR web site
            (http://www.tigr.org/tdb/at/at.html).

            Genes were identified by a combination of three methods: Gene
            prediction programs including GRAIL
            (ftp://arthur.epm.ornl.gov/pub/xgrail), Genefinder (Phil Green,
            University of Washington), Genscan (Chris Burge,
            http://gnomic.stanford.edu/GENSCANW.html), and NetPlantGene
            (http://www.cbs.dtu.dk/services/NetGene2/), searches of the
            complete sequence against a peptide database and plant EST
            databases at TIGR, and manual curations based on those analyses.
            Annotated genes are named to indicate the level of evidence for
            their annotation. Genes with similarity to other proteins are named
            after the database hits. Genes without significant peptide
            similarity but with EST similarity are named as 'unknown' proteins.
            Genes without protein or EST similarity, that are predicted by two
            or more gene prediction programs over most of their length are
            annotated as 'hypothetical' proteins. Genes encoding tRNAs are
            predicted by tRNAscan-SE (Sean Eddy,
            http://genome.wustl.edu/eddy/tRNAscan-SE/). Simple repeats were
            identified by repeatmasker (Arian Smit,
            http://ftp.genome.washington.edu/RM/RepeatMasker.html). Genes are
            numbered from the top to bottom of the chromosome.

            We thank the CSHL/WashU/ABI consortium for sequencing BAC clones
            F6P23, F5J6, T17A5, and T13L16, the ESSA group for sequencing clone
            F13D4, and Scott Jackson, Jiming Jiang, Klaus Meyer, Eric Richards
```

and Satoshi Tabata for helpful assistance. In addition, we would like to thank the TIGR Bioinformatics Department, especially Lixin Zhou, Hanif Khalak, Michael E. Heaney, Lily Fu, Feng Liang, Jeremy Peterson, Michael Holmes, and Delwood Richardson for software and database support.

This work was supported by the National Science Foundation, Department of Energy and the US Department of Agriculture.

Address all correspondence to: at@tigr.org.

FEATURES Location/Qualifiers

source 1. .95137
/organism="Arabidopsis thaliana"
/cultivar="Columbia"
/db_xref="taxon:3702"
/chromosome="II"

misc_feature <1. .300
/note="Sequence from clone F10A8"

misc_feature 301. .>95137
/note="Sequence from clone F2I9"

mRNA <341. .>997
/gene="At2g01390"

gene <341. .>997
/gene="At2g01390"
/note="F2I9.1; predicted by genscan"

CDS 341. .997
/gene="At2g01390"
/note="hypothetical protein"
/codon_start=1
/protein_id="AAC67316.1"
/db_xref="GI:3785969"
/translation="MKQNLVAVDILLNQMRDRNIKLDSFVVSAIIETNCDRCRTEGAS
LAFDYSLEMGIHLKKSAYLALIGNFLRSNELPKVIEVVKEMVKAQHSLGCYQGAMLIH
RLGFGRRPRLAADVFDLLPDDQKGVAAYTALMDVYISAGSPEKAMKILREMRREREIMP
SLGTYDVLLSGLEKTSDFQKEVALLRKEKKSLVASARFRENVHVEDKICDLLFATNLL
"

repeat_region complement(1085. .1127)
/rpt_family="(TAAAA)n"

mRNA join(<1175. .1294,1620. .>1940)
/gene="At2g01400"

gene <1175. .>1940
/gene="At2g01400"
/note="F2I9.2; predicted by genscan"

CDS join(1175. .1294,1620. .1940)
/gene="At2g01400"
/note="hypothetical protein"
/codon_start=1
/protein_id="AAC67317.1"
/db_xref="GI:3785970"
/translation="MARSCSIWTPVLISLSPVTGESPKISRRRVILATSIGSPQPLLE
AKIKEPQKLQVLDSKDVSRRNTMLYLAAGFLGGINFLNGEAAEARVGRKENRKKALEK
LRAKAKESEPNNKSGNQKIEKELEKEEVFPLLPPPLVVEANLLQ"

mRNA complement(<2069. .>3232)
/gene="At2g01410"

gene complement(<2069. .>3232)
/gene="At2g01410"
/note="F2I9.3; predicted by genefinder"

CDS complement(2069. .3232)
/gene="At2g01410"
/note="hypothetical protein"
/codon_start=1
/protein_id="AAC67318.1"
/db_xref="GI:3785971"
/translation="MKANMSSSSSTRYRSFFVFPILFAVLFLGLLIPSSADNRHVINF
RSPGLYPEGLTWDPRDQHFLVGSLHSRTIHSVSDAGVIETLISDLDLPENSTILGLAV
DSTNRRLLACIQSLPPLPPFSALASYDLRSGGRRVFLSPLPSLPGDDEDIARDVANDV
AVDFKGNAYVTNSAKNFIWKVDRDGAASIFSKSPLFNSQPVAADADASFRDCGLNGIV
YISKGYLLVVQSNTGKVFKVDEDSGNARLVLLNGDLIAADGMTRRRRDGTVMVVSQKK
LWLLKSQDSWSEGVVYDEIDLDIEGFPTAVTVAGRDRIYVLYGRVMEGIMGSSYKEEG
AREWFGIEEVWSEKEGGEDKIWLYVLIGFGFAYFCFWRFQMKKLITNMDKKIT"

mRNA complement(join(6967. .7404,7509. .7585,7680. .7837,
8215. .8300,8440. .8707,8865. .>10059))
/gene="At2g01420"

gene complement(6967. .>10059)
/gene="At2g01420"
/note="F2I9.4"

CDS complement(join(7338. .7404,7509. .7585,7680. .7837,
8215. .8300,8440. .8707,8865. .10059))
/gene="At2g01420"
/codon_start=1
/product="putative auxin transport protein"
/protein_id="AAC67319.1"
/db_xref="GI:3785972"
/translation="MITWHDLYTVLTAVVPLYVAMILAYGSVQWWKIFSPDQCSGINR
FVAIFAVPLLSFHFISTNDPYAMNFRFVAADTLQKIIMLVLLALWANLTKNGSLEWMI
TIFSLSTLPNTLVMGIPLLIAMYGTYAGSLMVQVVVLQCIIWYTLLLFLFEYRGAKLL
IMEQFPETGASIVSFKVESDVVSLDGHDFLETDAEIGNDGKLHVTVRKSNASRRSLMM
TPRPSNLTGAEIYSLSSTPRGSNFNHSDFYSVMGFPGGRLSNFGPADLYSVQSSRGPT
PRPSNFEENNAVKYGFYNNTNSSVPAAGSYPAPNPEFSTGTGVSTKPNKIPKENQQQL
QEKDSKASHDAKELHMFVWSSSASPVSDVFGGGAGDNVATEQSEQGAKEIRMVVSDQP
RKSNARGGGDDIGGLDSGEGEREIEKATAGLNKMGSNSTAELEAAGGDGGGNNGTHMP
PTSVMTRLILIMVWRKLIRNPNTYSSLIGLIWALVAYRWHVAMPKILQQSISILSDAG
LGMAMFSLGLFMALQPKIIACGNSVATFAMAVRFITGPAIMAVAGIAIGLHGDLLRIA
IVQAALPQGIVPFVFAKEYNVHPTILSTGVIFGMLIALPITLVYYILLGL"

repeat_region 14152. .14264
/rpt_family="POLY_A"

mRNA complement(join(<14658. .14916,16258. .16337,16458. .>16607))
/gene="At2g01430"

gene complement(<14658. .>16607)
/gene="At2g01430"
/note="F2I9.5"

CDS complement(join(14658. .14916,16258. .16337,16458. .16607))
/gene="At2g01430"
/codon_start=1
/product="putative homeodomain transcription factor"
/protein_id="AAC67320.1"
/db_xref="GI:3785973"
/translation="MNRLPSSEDGDDEEFSHDDGSAPPRKKLRLTREQSRLLEDSFRQ
NHTLNPKQKEVLAKHLMLRPRQIEVWFQNRRARSKLKQTEMECEYLKRWFGSLTEENH
RLHREVEELRAMKVGPTTVNSASSLTMCPRCERVTPAASPSRAVVPVPAKKTFPPQER
DR"

mRNA complement(join(<20954. .21559,21680. .21763,21869. .22027,
22234. .22377,22526. .22606,22742. .22813,23017. .23250,
23352. .23441,23559. .23640,23906. .24069,24308. .24542,
24746. .24858,24942. .>25415))
/gene="At2g01440"

gene complement(<20954. .>25415)
/gene="At2g01440"
/note="F2I9.6"

CDS complement(join(20954. .21559,21680. .21763,21869. .22027,
22234. .22377,22526. .22606,22742. .22813,23017. .23250,
23352. .23441,23559. .23640,23906. .24069,24308. .24542,
24746. .24858,24942. .25415))
/gene="At2g01440"
/codon_start=1
/product="putative ATP-dependent DNA helicase (RECG)"
/protein_id="AAC67321.1"
/db_xref="GI:3785974"
/translation="MECDNVDDFIDKKSDEQVKKDLVLACKRFPSIILGDSRPVELYS
NSKSYGESSSILKTPTDNSFLPTPMHGGWFDPDNLSRTLSSFCPELLQNDDSSDPRED
ILDDGSSFTSKTATSEVEATSDDVFAAQRFLATSIDSMPGLSKRHSNQLDSCGFHTMK
KLLHHFPRTYADLQNAQVDIEDGQYLIFVGKVLSSKGVRASSSFSFLEVIVSCEVSGR
DRTPEDLSHNAEDKAGKSIFLHLKKFFRGTRFTWQPFLNSIQEKHKVGDLVCISGKVK
SLRAEDHFEMREYNIDVLKDEEESSHRAQGRPYPIYPSKGGLNPKFLSDVISRALRVL
PANMDPIPKEITKVFGLPSLNDAYVGIHEPKTLDEADLARKRLIFDEFFYLQLARLYQ
MLQSLGTKIEKDVLLEKFRKPVLNSVYIEEWSTLTKSFLKALPYSLTPSQLSAVSEII
WDLKRPVPMNRLLQGDVGCGKTVVAFLACMEVIGSGYQAAFMAPTELLAIQHYEQCRD
LLENMEGDLQSGAISFIIGTHSLIAEKIEYSALRIAVVDEQQRFGVIQRGKFNSKLYG
TSMISKSGSSDSDDTSKADLSMAPHVLAMSATPIPRSLALALYGDISLTQITGMPLGR
IPVETHIFEGNETGIKEVYSMMLEDLKSGGRVYVVYPVIDQSEQLPQLRAASAELEIV

Query Match 8.1%; Score 140; DB 12; Length 95137;
Best Local Similarity 52.7%; Pred. No. 9.5e-20;
Matches 452; Conservative 0; Mismatches 385; Indels 21; Gaps 6;

Qy 560 cctttcataaggtttactcaattaaccgctaatcaagcgattttagaagcgattaacggt 619

```
Db  83307 CCTTATCTTAAATTCGCTCACTTCACGGCGAACCAAGCGATTCTCGAAGCTTTTGAAGGT 83248

Qy    620 aatcatcaagcaatccacatcgttgatttcgacattaatcacggggttcaatggccaccg 679
Db  83247 AAGAAGAGAG---TACACGTCATTGATTTCTCGATGAACCAAGGTCTTCAATGGCCTGCA 83191

Qy    680 ttaatgcaagcactagctgatcgtta---ccctgctcccactcttcgaatcaccggtact 736
Db  83190 CTTATGCAAGCTCTTGCGCTTCGAGAAGGAGGTCCTCCAACTTTCCGGTTAACCGGAATT 83131

Qy    737 gg---aaatgaccttgataccttcgtagaacaggtgatcgtttagctaaatttgctcac 793
Db  83130 GGTCCACCGGCGCCGGATAATTCTGATCATCTTCATGAAGTTGGTTGTAAATTAGCTCAG 83071

Qy    794 tcattagggttgagatttcaattccatcctctttatatagccaataataaccacgatcac 853
Db  83070 CTTGCGGAGGCGATTCACGTAGAATTCGAATACCGTGGATTCGTTGCTAACAGCTTAGCC 83011

Qy    854 gatgaagatccttctattatttcctccattgtactactccctgatgaaaccctagctatc 913
Db  83010 GATCTCGATGCTTCGATGCTTGAGC---TTAGACCGAGCGATACGGAAGCTGTTGCGGTG 82954

Qy    914 aactgtgttttctacctccaccgccttttaaaagaccgcgaaaagttaaggatttttttg 973
Db  82953 AACTCTGTTTTTGAGCTACATAAGCTCTTAGGTCGTCCCGGTGGGATAGAGAAAGTTCTC 82894

Qy    974 catagggttaagtcaatgaaccctaaaattgttacaatcgcggagaaggaagcaaatcat 1033
Db  82893 GGCGTTGTGAAACAGATTAAACCGGTGATTTTCACGGTGGTTGAGCAAGAATCGAACCAT 82834

Qy   1034 aaccatcctctttttttacaaagattcatcgaggcgttggattattatacagctgtgtttt 1093
Db  82833 AACGGACCGGTTTTCTTAGACCGGTTTACTGAATCGTTACATTATTATTCGACTCTGTTT 82774

Qy   1094 gattcactggaagctacattgccaccgggtagtcgagagaggatgacagttgaacaagtg 1153
Db  82773 GATTCGTTGGAAG------GAGTTCCGAATAGTCAAGACAAAGTCATGTCTG---AAGTT 82723

Qy   1154 tggtttgggagagagattgttgatatcgttgcgatggaaggagataaaaggaaagaaaga 1213
Db  82722 TACTTAGGGAAACAGATTTGTAATCTGGTGGCTTGTGAAGGTCCTGACAGAGTCGAGAGA 82663

Qy   1214 catgaaaggtttagatcatgggaagttatgttgaggagttgtggatttagtaatgttgct 1273
Db  82662 CACGAAACGTTGAGTCAATGGGGAAACCGGTTTGGTTCGTCCGGTTTAGCGCCGGCACAT 82603

Qy   1274 ttaagccctttttgcattatcacaagctaagcttctttttgagacttcattatccttctgaa 1333
Db  82602 CTTGGGTCTAACGCGTTTAAGCAAGCGAGTATGCTTTTGTCTGTGTTTAATAGTGGCCAA 82543

Qy   1334 ggctatcaactcggagtttcgagtaattctttcttcttaggttggcaaaatcaacccctt 1393
Db  82542 GGTTATCGTGTGGAGGAGAGTAATGGATGTTTGATGTTGGGTTGGCACACTCGTCCACTC 82483

Qy   1394 ttctccatctcgtcttgg 1411
Db  82482 ATTACCACCTCCGCTTGG 82465
```

```
RESULT  12
ATRGA2
LOCUS       ATRGA2      1951 bp    mRNA            PLN       19-AUG-1997
DEFINITION  A.thaliana mRNA for RGA2 gene.
ACCESSION   Y11337
VERSION     Y11337.1  GI:2339979
KEYWORDS    RGA2 gene.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
            Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
            Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 1951)
  AUTHORS   Truong,H.N., Caboche,M. and Daniel-Vedele,F.
  TITLE     Sequence and characterization of two Arabidopsis thaliana cDNAs
            isolated by functional complementation of a yeast gln3 gdh1 mutant
  JOURNAL   FEBS Lett. 410 (2-3), 213-218 (1997)
  MEDLINE   97379310
REFERENCE   2  (bases 1 to 1951)
  AUTHORS   Truong,H.N.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-FEB-1997) H.N. Truong, INRA-Versailles, Laboratoire
            de Biologie Cellulaire, Route de Saint-Cyr, 78026 Versailles Cedex,
            FRANCE
FEATURES             Location/Qualifiers
     source          1. .1951
                     /organism="Arabidopsis thaliana"
                     /sub_species="(L). Heynh"
                     /db_xref="taxon:3702"
                     /dev_stage="two-leaf stage seedlings"
                     /note="Landsberg erecta ecotype"
     gene            196. .1794
                     /gene="RGA2"
     CDS             196. .1794
                     /gene="RGA2"
                     /codon_start=1
                     /product="RGA2 protein"
                     /protein_id="CAA72178.1"
                     /db_xref="GI:2339980"
                     /db_xref="SPTREMBL:O23643"
                     /translation="MKRDHHHHHQDKKTMMMNEEDDGNGMDELLAVLGYKVRSSEMAD
                     VAQKLEQLEVMMSNVQEDDLSQLATETVHYNPAELYTWLDSMLTDLNPPSSNAEYDLK
                     AIPGDAILNQFAIDSASSSNQGGGGDTYTTNKRLKCSNGVVETTTATAESTRHVVLVD
                     SQENGVRLVHALLACAEAVQKENLTVAEALVKQIGFLAVSQIGAMRQVATYFAEALAR
                     RIYRLSPSQSPIDHSLSDTLQMHFYETCPYLKFAHFTANQAILEAFQGKKRVHVIDFS
                     MSQGLQWPALMQALALRPGGPPVFRLTGIGPPAPDNFDYLHEVGCKLAHLAEAIHVEF
                     EYRGFVANTLADLDASMLELRPSEIESVAVNSVFELHKLLGRPGAIDKVLGVVNQIKP
                     EIFTVVEQESNHNSPIFLDRFTESLHYYSTLFDSLEGVPSGQDKVMSEVYLGKQICNV
                     VACDGPDRVERHETLSQWRNRFGSAGFAAAHIGSNAFKQASMLLALFNGGEGYRVEES
                     DGCLMLGWHTRPLIATSAWKLSTN"
BASE COUNT      482 a    419 c    475 g    575 t
ORIGIN
```

```
  Query Match             7.7%;  Score 133.6;  DB 13;  Length 1951;
  Best Local Similarity   52.2%;  Pred. No. 1.7e-18;
  Matches  448;  Conservative    0;  Mismatches  389;  Indels   21;  Gaps    6;

Qy    560 cctttcataaggtttactcaattaaccgctaatcaagcgattttagaagcgattaacggt 619
Db    934 CCTTATCTCAAGTTCGCTCACTTCACGGCGAATCAAGCGATTCTCGAAGCTTTTCAAGGG 993

Qy    620 aatcatcaagcaatccacatcgttgatttcgacattaatcacggggttcaatggccaccg 679
Db    994 AAGAAAAGAG---TTCATGTCATTGATTTCTCTATGAGTCAAGGTCTTCAATGGCCGGCG 1050

Qy    680 ttaatgcaagcactagctgatcgttaccctgctcccactctt---cgaatcaccggtact 736
Db   1051 CTTATGCAGGCTCTTGCGCTTCGACCTGGTGGTCCTCCTGTTTTCCGGTTAACCGGAATT 1110

Qy    737 gg---aaatgaccttgataccttcgtagaacaggtgatcgtttagctaaatttgctcac 793
Db   1111 GGTCCACCGGCACCGGATAATTTCGATTATCTTCATGAAGTTGGGTGTAAGCTGGCTCAT 1170

Qy    794 tcattagggttgagatttcaattccatcctctttatatagccaataataaccacgatcac 853
Db   1171 TTAGCTGAGGCGA---TTCACGTTGAGTTTGAGTACAGAGGATTTGTGGCTAACACTTTA 1227

Qy    854 gatgaagatccttctattatttcctccattgtactactccctgatgaaaccctagctatc 913
Db   1228 GCTGATCTTGATGCTTCGATGCTTGAGCTTAGACCAAGTGAGATTGAATCTGTTGCGGTT 1287

Qy    914 aactgtgttttctacctccaccgccttttaaaagaccgcgaaaagttaaggatttttttg 973
Db   1288 AACTCTGTTTTCGAGCTTCACAAGCTCTTGGGACGACCTGGTGCGATCGATAAGGTTCTT 1347
```

```
Qy    974 catagggttaagtcaatgaaccctaaaattgttacaatcgcggagaaggaagcaaatcat 1033
          |  ||| ||    || || ||  | ||| | ||  | |  ||| ||||| | || |||
Db   1348 GGTGTGGTGAATCAGATTAAACCGGAGATTTTCACTGTGGTTGAGCAGGAATCGAACCAT 1407

Qy   1034 aaccatcctctttttttacaaagattcatcgaggcgttggattattatacagctgtgttt 1093
          ||   |||  |||| ||| |  | || |  ||| ||||| |||||||  |  |  |||||
Db   1408 AATAGTCCGATTTTCTTAGATCGGTTTACTGAGTCGTTGCATTATTACTCGACGTTGTTT 1467

Qy   1094 gattcactggaagctacattgccaccgggtagtcgagagaggatgacagttgaacaagtg 1153
          || ||  |||||| |        |||| || ||| ||| | | | |    |      | ||
Db   1468 GACTCGTTGGAAGGT------GTACCGAGTGGTCAAGACAAGGTCA---TGTCGGAGGTT 1518

Qy   1214 catgaaaggtttagatcatgggaagttatgttgaggagttgtggatttagtaatgttgct 1273
          ||||||| ||| ||    |||        ||| ||  |  ||| |||    ||    |
Db   1579 CATGAAACGTTGAGTCAGTGGAGGAACCGGTTCGGGTCTGCTGGGTTTGCGGCTGCACAT 1638

Qy   1274 ttaagccctttttgcattatcacaagctaagcttcttttgagacttcattatccttctgaa 1333
          |  |  |  ||| ||    ||||| |   | ||||||   ||    |      |||
Db   1639 ATTGGTTCGAATGCGTTTAAGCAAGCGAGTATGCTTTTGGCTCTGTTCAACGGCGGTGAG 1698

Qy   1334 ggctatcaactcggagtttcgagtaattctttcttcttaggttggcaaaatcaacccctt 1393
          || ||||   | |  |           | | || | || |||||||| |  |  ||| ||
Db   1699 GGTTATCGGGTGGAGGAGAGTGACGGCTGTCTCATGTTGGGTTGGCACACACGACCGCTC 1758

Qy   1394 ttctccatctcgtcttgg 1411
            |  ||| |||| |||||
Db   1759 ATAGCCACCTCGGCTTGG 1776
```

```
RESULT  13
A64697
LOCUS       A64697       1964 bp    DNA             PAT       29-MAR-1999
DEFINITION  Sequence 1 from Patent WO9729123.
ACCESSION   A64697
VERSION     A64697.1  GI:4530762
KEYWORDS    .
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
            Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
            Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 1964)
  AUTHORS   Harberd,N.P., Peng,J., Carol,P. and Richards,D.E.
  TITLE     NUCLEIC ACID ENCODING GAI GENE OF ARABIDOPSIS THALIANA
  JOURNAL   Patent: WO 9729123-A 1 14-AUG-1997;
            INNES JOHN CENTRE INNOV LTD (GB)
COMMENT     Other publication AU 1799697 19970828.
FEATURES             Location/Qualifiers
     source          1. .1964
                     /organism="Arabidopsis thaliana"
                     /strain="LANDSBERG ERECTA"
                     /db_xref="taxon:3702"
                     /chromosome="1"
                     /map="20.9"
                     /haplotype="GAI"
BASE COUNT      489 a    426 c    474 g    575 t
ORIGIN
```

```
  Query Match             7.7%;  Score 133.6;  DB 9;  Length 1964;
  Best Local Similarity   52.2%;  Pred. No. 1.7e-18;
  Matches  448;  Conservative    0;  Mismatches  389;  Indels   21;  Gaps    6;

Qy    560 cctttcataaggtttactcaattaaccgctaatcaagcgattttagaagcgattaacggt 619
          ||||   | | |||  |||| || || || ||||||||||||| | |||||  || | ||
Db    947 CCTTATCTCAAGTTCGCTCACTTCACGGCGAATCAAGCGATTCTCGAAGCTTTTCAAGGG 1006

Qy    620 aatcatcaagcaatccacatcgttgatttcgacattaatcacggggttcaatggccaccg 679
          ||  |   ||   | ||  || ||||||||   || | ||| ||  ||||||||||  ||
```

```
Db   1007 AAGAAAAGAG---TTCATGTCATTGATTTCTCTATGAGTCAAGGTCTTCAATGGCCGGCG 1063

Qy    680 ttaatgcaagcactagctgatcgttaccctgctcccactctt---cgaatcaccggtact 736
          | |||||| || || ||   |||       || |||  || ||   ||  | |||||| | |
Db   1064 CTTATGCAGGCTCTTGCGCTTCGACCTGGTGGTCCTCCTGTTTTCCGGTTAACCGGAATT 1123

Qy    737 gg---aaatgaccttgatacccttcgtagaacaggtgatcgtttagctaaatttgctcac 793
          ||   |   |  |||| |  |        |||  |     |||  | ||||||
Db   1124 GGTCCACCGGCACCGGATAATTTCGATTATCTTCATGAAGTTGGGTGTAAGCTGGCTCAT 1183

Qy    794 tcattagggttgagatttcaattccatcctctttatatagccaataataaccacgatcac 853
          | |   | |  ||  |||| |  |   | || ||    |       ||  |
Db   1184 TTAGCTGAGGCGA---TTCACGTTGAGTTTGAGTACAGAGGATTTGTGGCTAACACTTTA 1240

Qy    854 gatgaagatccttctattatttcctccattgtactactccctgatgaaaccctagctatc 913
          | |||   | | ||   ||         ||  || |      ||||  |  | ||  |
Db   1241 GCTGATCTTGATGCTTCGATGCTTGAGCTTAGACCAAGTGAGATTGAATCTGTTGCGGTT 1300

Qy    914 aactgtgtttttctacctccaccgccttttaaaagaccgcgaaaagttaaggatttttttg 973
          |||| |||||||  | || |||   || ||   |  | |  |  |  |    || |
Db   1301 AACTCTGTTTTCGAGCTTCACAAGCTCTTGGGACGACCTGGTGCGATCGATAAGGTTCTT 1360

Qy    974 catagggttaagtcaatgaaccctaaaattgttacaatcgcggagaaggaagcaaatcat 1033
            |  ||| ||    || || ||  | ||| | ||  | |  ||| ||||| | || |||
Db   1361 GGTGTGGTGAATCAGATTAAACCGGAGATTTTCACTGTGGTTGAGCAGGAATCGAACCAT 1420

Qy   1034 aaccatcctctttttttacaaagattcatcgaggcgttggattattatacagctgtgttt 1093
          ||   |||  |||| ||| |  | || |  ||| ||||| |||||||  |  |  |||||
Db   1421 AATAGTCCGATTTTCTTAGATCGGTTTACTGAGTCGTTGCATTATTACTCGACGTTGTTT 1480

Qy   1094 gattcactggaagctacattgccaccgggtagtcgagagaggatgacagttgaacaagtg 1153
          || ||  |||||| |        |||| || ||| ||| | | | |    |      | ||
Db   1481 GACTCGTTGGAAGGT------GTACCGAGTGGTCAAGACAAGGTCA---TGTCGGAGGTT 1531

Qy   1154 tggtttgggagagagattgttgatatcgttgcgatggaaggagataaaaggaaagaaaga 1213
          |  || || | | ||||    |  | || ||    || ||| | |  |   || |
Db   1532 TACTTGGGTAAACAGATCTGCAACGTTGTGGCTTGTGATGGACCTGACCGAGTTGAGCGT 1591

Qy   1214 catgaaaggtttagatcatgggaagttatgttgaggagttgtggatttagtaatgttgct 1273
          ||||||| ||| ||    |||        ||| ||  |  ||| |||    ||    |
Db   1592 CATGAAACGTTGAGTCAGTGGAGGAACCGGTTCGGGTCTGCTGGGTTTGCGGCTGCACAT 1651

Qy   1274 ttaagccctttttgcattatcacaagctaagcttcttttgagacttcattatccttctgaa 1333
          |  |  |  ||| ||    ||||| |   | ||||||   ||    |      |||
Db   1652 ATTGGTTCGAATGCGTTTAAGCAAGCGAGTATGCTTTTGGCTCTGTTCAACGGCGGTGAG 1711

Qy   1334 ggctatcaactcggagtttcgagtaattctttcttcttaggttggcaaaatcaacccctt 1393
          || ||||   | |  |           | | || | || |||||||| |  |  ||| ||
Db   1712 GGTTATCGGGTGGAGGAGAGTGACGGCTGTCTCATGTTGGGTTGGCACACACGACCGCTC 1771

Qy   1394 ttctccatctcgtcttgg 1411
            |  ||| |||| |||||
Db   1772 ATAGCCACCTCGGCTTGG 1789
```

```
RESULT  14
ATY15193
LOCUS       ATY15193     1964 bp    DNA             PLN       01-NOV-1997
DEFINITION  Arabidopsis thaliana GAI gene.
ACCESSION   Y15193
VERSION     Y15193.1  GI:2569937
KEYWORDS    GAI gene.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
            Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
            Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 1964)
  AUTHORS   Peng,J., Carol,P., Richards,D.E., King,K.E., Cowling,R.J.,
            Murphy,G.P. and Harberd,N.P.
  TITLE     The Arabidopsis GAI gene defines a signalling pathway that
            negatively regulates gibberellin responses
  JOURNAL   Genes Dev. In press
```

```
REFERENCE   2  (bases 1 to 1964)
  AUTHORS   Harberd,N.P.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-OCT-1997) N.P. Harberd, John Innes Centre, Molecular
            Genetics, Colney Lane, Norwich, NR4 7UJ, UK
FEATURES             Location/Qualifiers
     source          1. .1964
                     /organism="Arabidopsis thaliana"
                     /strain="Landsberg erecta"
                     /db_xref="taxon:3702"
                     /chromosome="1"
     gene            209. .1807
                     /gene="GAI"
     CDS             209. .1807
                     /gene="GAI"
                     /codon_start=1
                     /protein_id="CAA75492.1"
                     /db_xref="GI:2569938"
                     /db_xref="SPTREMBL:O23724"
                     /translation="MKRDHHHHHQDKKTMMMNEEDDGNGMDELLAVLGYKVRSSEMAD
                     VAQKLEQLEVMMSNVQEDDLSQLATETVHYNPAELYTWLDSMLTDLNPPSSNAEYDLK
                     AIPGDAILNQFAIDSASSSNQGGGGDTYTTNKRLKCSNGVVETTTATAESTRHVVLVD
                     SQENGVRLVHALLACAEAVQKENLTVAEALVKQIGFLAVSQIGAMRKVATYFAEALAR
                     RIYRLSPSQSPIDHSLSDTLQMHFYETCPYLKFAHFTANQAILEAFQGKKRVHVIDFS
                     MSQGLQWPALMQALALRPGGPPVFRLTGIGPPAPDNFDYLHEVGCKLAHLAEAIHVEF
                     EYRGFVANTLADLDASMLELRPSEIESVAVNSVFELHKLLGRPGAIDKVLGVVNQIKP
                     EIFTVVEQESNHNSPIFLDRFTESLHYYSTLFDSLEGVPSGQDKVMSEVYLGKQICNV
                     VACDGPDRVERHETLSQWRNRFGSAGFAAAHIGSNAFKQASMLLALFNGGEGYRVEES
                     DGCLMLGWHTRPLIATSAWKLSTN"
BASE COUNT      489 a    426 c    474 g    575 t
ORIGIN
```

```
  Query Match             7.7%;  Score 133.6;  DB 14;  Length 1964;
  Best Local Similarity   52.2%;  Pred. No. 1.7e-18;
  Matches  448;  Conservative    0;  Mismatches  389;  Indels   21;  Gaps    6;

Qy     560 cctttcataaggtttactcaattaaccgctaatcaagcgattttagaagcgattaacggt 619
           ||||   | | |||  |||| || || || ||||||||||||| | |||||  || | ||
Db     947 CCTTATCTCAAGTTCGCTCACTTCACGGCGAATCAAGCGATTCTCGAAGCTTTTCAAGGG 1006

Qy     620 aatcatcaagcaatccacatcgttgatttcgacattaatcacggggttcaatggccaccg 679
           ||  |   ||   | ||  || ||||||||   || | ||| ||  ||||||||||  ||
Db    1007 AAGAAAAGAG---TTCATGTCATTGATTTCTCTATGAGTCAAGGTCTTCAATGGCCGGCG 1063

Qy     680 ttaatgcaagcactagctgatcgttaccctgctcccactctt---cgaatcaccggtact 736
            | ||||| || || ||   |||       || |||  || ||   ||  | ||||| | |
Db    1064 CTTATGCAGGCTCTTGCGCTTCGACCTGGTGGTCCTCCTGTTTTCCGGTTAACCGGAATT 1123

Qy     737 gg---aaatgaccttgataccttcgtagaacaggtgatcgtttagctaaatttgctcac 793
           ||   |   | |  ||||   |  |        |||   |     |||  | |||||
Db    1124 GGTCCACCGGCACCGGATAATTTCGATTATCTTCATGAAGTTGGGTGTAAGCTGGCTCAT 1183

Qy     794 tcattagggttgagatttcaattccatcctctttatatagccaataataaccacgatcac 853
           | |   | |  ||   ||||  |  |   |   || | ||    |       ||  |
Db    1184 TTAGCTGAGGCGA---TTCACGTTGAGTTTGAGTACAGAGGATTTGTGGCTAACACTTTA 1240

Qy     854 gatgaagatccttctattatttcctccattgtactactccctgatgaaaccctagctatc 913
           | |||   |  | ||   ||         ||  || |        |||| |  | ||  |
Db    1241 GCTGATCTTGATGCTTCGATGCTTGAGCTTAGACCAAGTGAGATTGAATCTGTTGCGGTT 1300

Qy     914 aactgtgttttctacctccaccgccttttaaaagaccgcgaaaagttaaggatttttttg 973
           |||| |||||| | || |||   || ||   |  |  |  |   |  || |
Db    1301 AACTCTGTTTTCGAGCTTCACAAGCTCTTGGGACGACCTGGTGCGATCGATAAGGTTCTT 1360

Qy     974 catagggttaagtcaatgaaccctaaaattgttacaatcgcggagaaggaagcaaatcat 1033
             | ||| ||   || || ||  | ||| | ||  | |  ||| ||||| | || |||
Db    1361 GGTGTGGTGAATCAGATTAAACCGGAGATTTTCACTGTGGTTGAGCAGGAATCGAACCAT 1420

Qy    1034 aaccatcctctttttttacaaagattcatcgaggcgttggattattatacagctgtgttt 1093
           ||  |||  |||| ||| |  | || |  ||| ||||| |||||||  |  |  |||||
Db    1421 AATAGTCCGATTTTCTTAGATCGGTTTACTGAGTCGTTGCATTATTACTCGACGTTGTTT 1480

Qy    1094 gattcactggaagctacattgccaccgggtagtcgagagaggatgacagttgaacaagtg 1153
           || ||  |||||| |        |||| || ||| ||| | | | |   |     | ||
Db    1481 GACTCGTTGGAAGGT------GTACCGAGTGGTCAAGACAAGGTCA---TGTCGGAGGTT 1531

Qy    1154 tggtttgggagagagattgttgatatcgttgcgatggaaggagataaaaggaaagaaaga 1213
           |  || || | | ||||     |  | || ||    || |||  | |  |    ||  |
Db    1532 TACTTGGGTAAACAGATCTGCAACGTTGTGGCTTGTGATGGACCTGACCGAGTTGAGCGT 1591

Qy    1214 catgaaaggtttagatcatgggaagttatgttgaggagttgtggatttagtaatgttgct 1273
           ||||||| ||| ||    |||        |||  ||  |  ||| |||     ||    |
Db    1592 CATGAAACGTTGAGTCAGTGGAGGAACCGGTTCGGGTCTGCTGGGTTTGCGGCTGCACAT 1651

Qy    1274 ttaagccctttttgcattatcacaagctaagcttcttttgagacttcattatccttctgaa 1333
            |  |  |   ||| ||    ||||| |   | ||||||   ||     |      |||
Db    1652 ATTGGTTCGAATGCGTTTAAGCAAGCGAGTATGCTTTTGGCTCTGTTCAACGGCGGTGAG 1711

Qy    1334 ggctatcaactcggagtttcgagtaattctttcttcttaggttggcaaaatcaacccctt 1393
           || ||||   | |  |           | | || | || |||||||| |  | ||| ||
Db    1712 GGTTATCGGGTGGAGGAGAGTGACGGCTGTCTCATGTTGGGTTGGCACACACGACCGCTC 1771

Qy    1394 ttctccatctcgtcttgg 1411
            |  ||| |||| |||||
Db    1772 ATAGCCACCTCGGCTTGG 1789
```

```
RESULT  15
AC006917
LOCUS       AC006917   132699 bp    DNA             PLN       28-JUN-2000
DEFINITION  Genomic sequence for Arabidopsis thaliana BAC F10B6 from chromosome
            I, complete sequence.
ACCESSION   AC006917
VERSION     AC006917.6  GI:4757662
KEYWORDS    HTG.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
            Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
            Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 132699)
  AUTHORS   Chao,Q., Shinn,P., Dunn,P., Buehler,E., Kahn,S., Kim,C., Walker,M.,
            Williams,S., Altafi,H., Araujo,R., Conn,L., Conway,A.B.,
            Gonzalez,A., Hansen,N.F., Huizar,L., Kremenetskaia,I., Lenz,C.,
            Li,J., Liu,S., Luros,S., Rowley,D., Schwartz,J., Toriumi,M.,
            Vysotskaia,V., Yu,G., Davis,R.W., Federspiel,N.A., Theologis,A. and
            Ecker,J.R.
  TITLE     Genomic sequence for Arabidopsis thaliana BAC F10B6 from chromosome
            I
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 132699)
  AUTHORS   Ecker,J.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-FEB-1999) Arabidopsis thaliana Genome Center,
            Department of Biology, University of Pennsylvania, 38th Street and
            Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA
REFERENCE   3  (bases 1 to 132699)
  AUTHORS   Ecker,J.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-MAY-1999) Arabidopsis thaliana Genome Center,
            Department of Biology, University of Pennsylvania, 38th Street and
            Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA
REFERENCE   4  (bases 1 to 132699)
  AUTHORS   Cheuk,R., Shinn,P., Brooks,S., Buehler,E., Chao,Q.,
            Johnson-Hopson,C., Khan,S., Kim,C., Altafi,H., Bei,B., Chin,C.,
            Chiou,J., Choi,E., Conn,L., Conway,A., Gonzalez,A., Hansen,N.,
            Howing,B., Koo,T., Lam,B., Lee,J., Lenz,C., Li,J., Liu,A., Liu,J.,
            Liu,S., Mukharsky,N., Nguyen,M., Palm,C., Pham,P., Sakano,H.,
            Schwartz,J., Southwick,A., Thaveri,A., Toriumi,M., Vaysberg,M.,
            Yu,G., Davis,R., Federspiel,N., Theologis,A. and Ecker,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-JUN-2000) Arabidopsis thaliana Genome Center,
            Department of Biology, University of Pennsylvania, 38th and
            Hamilton Walk, Philadelphia, PA 19104-6018, USA
COMMENT     On May 7, 1999 this sequence version replaced gi:4731042.
```

```
FEATURES             Location/Qualifiers
     source          1. .132699
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
                     /chromosome="1"
     CDS             complement(join(144. .278,360. .491,576. .673,814. .1435,
                     1749. .1867,2161. .2310,2441. .2565,2698. .2742,2837. .2902,
                     2979. .3031,3117. .3234,3314. .3415,3498. .3655,3741. .3818,
                     3894. .3985,4054. .4104,4138. .4420))
                     /note="similar to Na/H antiporter dbj|BAA89487.1"
                     /codon_start=1
                     /evidence=not_experimental
                     /product="F10B6.1"
                     /protein_id="AAF79251.1"
                     /db_xref="GI:8778242"
                     /translation="MTSIIGAALPYKSPEKAIASSSYSAENDSSPVDAVIFAGTSLVL
                     GTACRYLFNGTRVPYTVVLLVIGIFLGSLGLICLLIKLYSSVSISWKILEKYGTKHNL
                     GKLGHGIRIWNGINPDLLLAVFLPVLLFESSFSMDVHQIKRCMGQMVLLAGPGVLIST
                     FCLGALIKLTFPYNWDWKTSLLLGGLLGATDPVAVVALLKELGASKKMTTLIDGESLM
                     NDGVSVVVFQLFFKMVMGHNSDWGSIIKFLVQNSFGAVGIGLAFGIASVFWLKFIFND
                     TVAQITVTLSASYFAYYTAQEWAGVSGILTVMILGMFFAAFARTAFKGDSHQSLHHFW
                     EMAAYIANTLVFMLSGVIIAESVLSGQTISYKGVSSQFIILRPDFFFILFLEDVKGGN
                     SWSFLFLLYLYVQLSRCVVVGVLYPLLCRSGYGLDWKESIILTWSGLRAVLQFLFLTG
                     GIVFLTLVVNGSTTQLLLHLLRMDTLTATKKRILEYTKFEMMKTALKAFENLGDDEEL
                     GSADWPTVIRHISSLKDLEGRQVNPHDGYEAGSLDPTNIMDIRIRFLNGISQCVVNSS
                     HCQTSVVPKVAIDSRLLPVSGVQAAYWEMLDDGRITQCTANVLMQSVDEALDLVSTSS
                     LSDWRGLEPRVHFPNYYKFLQSKIIPHKLVTHLIVERLESACYISSAFLRAHRIARQQ
                     LHIFLGNSNIASTVINESEVEGEEAKQFLEDVRDSFPQVLSVLKTRQVTHYVLNHLNG
                     YIKNLEKVGLLEGKEVSHLHDVVQSDLKKLLRHPPSLKLPNVDDLITSNPLLKDRSSF
                     RSLAIGETDA"
     CDS             join(5650. .5801,6081. .6121,6230. .6819)
                     /note="unknown protein;  similar to EST emb|F14394.1"
                     /codon_start=1
                     /evidence=not_experimental
                     /product="F10B6.2"
                     /protein_id="AAF79216.1"
                     /db_xref="GI:8778207"
                     /translation="MRTPTTILLLVGAILFSGAGYVRSDASDHRYKEGDTVPLYANKV
                     GPFHNPSETYRYFDLPFCIPEGVKEKKEALGEVLNGDRLVSAPYKLNFRDEKESEVYC
                     NKKLSKEEVKQFRKAVEKDYYFQMYYDDLPIWGFIGKVDKDIKSDPSEFKYFLYKHIQ
                     FEILYNKDRVIEISARMDPHSLVDLTEDKEVDAEFMYTVKWKETETPFEKRMEKYSMS
                     SSLPHHLEIHWFSIINSCVTVLLLTGFLATILMRVLKNDFMK"
     CDS             join(6881. .7029,7140. .7296,7392. .7876,7961. .8180)
                     /note="similar to transmembrane 9 superfamily member 2
                     gi|4758874;  similar to dbj|AV442489.1, gb|AI996021.1,
                     emb|F14395.1"
                     /codon_start=1
                     /evidence=not_experimental
                     /product="F10B6.3"
                     /protein_id="AAF79217.1"
                     /db_xref="GI:8778208"
                     /translation="MLKRYAQDEEAADDQEETGWKYIHGDVFRFPTHNSLFAASLGSG
                     TQLFTLTIFIFMLALVGVFYPYNRGALFTALVVIYALTSGIAGYTSASFYCQLEGKSW
                     VRNLLLTGCLFCGPLFLTFCFLNTVAITYTATAALPFGTIVVIVLIWTLVTSPLLVLG
                     GIAGKNSKAEFQAPCRTTKYPREIPPLPWYRSAIPQMAMAGFLPFSAIYIELYYIFAS
                     VWGHRIYTIYSILFIVFIILIIVTAFITVALTYFQLAAEDHQWWWRSFLCGGSTGLFI
                     YAYCLYYYYARSDMSGFMQTSFFFGYMACICYGFFLMLGTVGFRAALLFVRHIYRSIK
                     CE"
     CDS             8842. .9714
                     /note="hypothetical protein"
                     /codon_start=1
                     /evidence=not_experimental
                     /product="F10B6.4"
                     /protein_id="AAF79218.1"
                     /db_xref="GI:8778209"
                     /translation="MDLSREVDDFIKETIDHSLGLPISMDVLQKKLYTAEESQRRLRE
                     QYLSLVSRLKEKEKVIDLVRSEASMNAQSLKKFVEENQKLGSEREDLVNQCKKWEKEC
                     FLYHQDRESLMEFGNETDERAREAESRVRELEEEVRKMSDEIKSRIESEEDCLVDSIL
                     ASFVSKDESISLGRIFLEANSEEDKCCETLLRKWDQLKPTTQKFVSLVSMVKRIEKEK
                     ECLIMNLAKAEQEVELVSEQNRELDRENRKFLRQCSAERSHGSNKFNKRKSIKMMSSP
                     IEKRIELSSQEFID"
     CDS             11067. .11906
                     /note="unknown protein;  similar to EST gb|AI995472.1"
                     /codon_start=1
                     /evidence=not_experimental
                     /product="F10B6.5"
                     /protein_id="AAF79219.1"
                     /db_xref="GI:8778210"
                     /translation="MDDDGFRNWGYYEPAAATFKGNLGLQLMSTIDRNTKPFLPGRDP
                     NLMMGPNGSYHHQEPPIHMSYNWINQQKDKFFNMLPVTTATPNYGNVLPETSSAPSMQ
                     MNLHHHLQTEENPVKLEEEIVVQTKKRKTNAKAGSTPKAKKPRKPKDENSNNNNNNNT
                     NVTRVKPAKKSVDLVINGVSMDISGLPVPICTCTGAPQQCYRWGCGGWQSACCTTNIS
                     MHPLPMSTKRRGARISGRKMSQGAFKKVLEKLASDGFNFGNPIDLKSHWARHGTNKFV
                     TIR"
     CDS             complement(13762. .14781)
                     /note="hypothetical protein"
                     /codon_start=1
                     /evidence=not_experimental
                     /product="F10B6.6"
                     /protein_id="AAF79250.1"
                     /db_xref="GI:8778241"
                     /translation="MKLWKRAAVVLKDGPSLIAADDILTAAVVKATSHDELSIDTESA
                     QFIYRHVLSSPSSLKPLVSLISSRVKRTRSWAVALKGLMLMHGFFLCKSTVAESIGRL
                     PFDLSSFGEGNSRIMSKSGGFNLFVRAYFAFLDRRSILFHDGNRHRYNEESSVLIRLV
                     IIRKMQIIVDSLIRIKPIGENMMIPVINEAMENVVSEIMEIYGWICRRIAEVLPNVHS
                     KIGKTEADLALKIVAKSMKQGGELKKYFEFCKDLGVSNAQEIPNFVRIPEADVIHLDE
                     LVRTAMESSEESAERTEIAEEEEEEEEEIETKLSDLITLDHNEEAPASPPRVVVVDIP
                     DLISF"
     CDS             16027. .16533
                     /note="hypothetical protein"
                     /codon_start=1
                     /evidence=not_experimental
                     /product="F10B6.7"
                     /protein_id="AAF79220.1"
                     /db_xref="GI:8778211"
                     /translation="MQSTCVYRECMRNHAAKLGSYAIDGCREYSQPSTGDLCVACGCH
                     RSYHRRIDVISSPQINHTRFPFTSLRRVKQLARLKWKTAEERNEEEEDDTEETSTEEK
                     MTVQRRRKSKFTAEQREAMKDYAAKLGWTLKDKRALREEIRVFCEGIGVTRYHFKTWV
                     NNNKKFYH"
     CDS             complement(join(17635. .17815,17901. .18019,18141. .18194,
                     18300. .18406,18547. .18611,18742. .18950))
                     /note="hypothetical protein"
                     /codon_start=1
                     /evidence=not_experimental
                     /product="F10B6.8"
                     /protein_id="AAF79249.1"
                     /db_xref="GI:8778240"
                     /translation="MEFKAYFILYWFQPFRLQCEPVIVLFLCFQMICLKKKQRIFLFW
                     YPEELHLQQPLTPKMVIFWVLFLRIRWWFDKQVCKSSSTTKGDSLVSVGAEGALGFNE
                     IVRPYEVKSLGESFKSLCSPKEGSKVIFGQALKVGTSLTFVGEAVRDKAGNLMIQKSK
                     EQSLIVFSEESSFDEMVNNMKSQSELCVILAKIFGSIAVAIAVVYGVDYARKVLLPFV
                     WKKKDLGDSSSQNGSSDSEDTDISDF"
     CDS             complement(join(20071. .20409,20499. .20573,20648. .20721,
                     20800. .20887,20957. .21108,21188. .21356,21448. .21906,
                     22006. .22275,22367. .22570,22840. .22917))
                     /note="hypothetical protein"
                     /codon_start=1
                     /evidence=not_experimental
                     /product="F10B6.9"
                     /protein_id="AAF79248.1"
                     /db_xref="GI:8778239"
                     /translation="MLEIESPTSLCFRTNTTCNALLRELQKIWVDIGESDAEKDRMLM
                     ELEKECLEIYRRKVDEAANSKAQLHQSLVSIEAEIASLLAALGVFNSHSPMKAKEGSK
                     SLKEKLAAVRPMLEDLRLQKDERMKQFVDIKAQIEKMSGEISGYSDQLNKTMVGSLAL
                     DEQDLTLRKLNEYQTHLRSLQKEKSDRLNKVLDYVNEVHTLCGVLGVDFGQTVSEVHP
                     SLHRTDHEQSTNISDDTLDGLHHMIHKLKTERSVRFQKVDKHACNAILLFVTDYHFDA
                     LAFSPLSVQLKDVAGSLFELWNLMDTSQEERTKFASVSYVVRSSESDITEPNILSSET
                     IEQVSAEVDCFNKLKASRMKELVMKRRTELENLCRLAHIEADTSTSLEKSTALIDSGT
                     KRLVDPSELLTNIELHINKIKEEAHSRKEIIDRIDRWLSACEEENWLEEYNQDETRYS
                     AGRGGHVNLKHAERARITVNKIPSMVDNLIKKTLLWEDETRKSFLYDGVRLVSILEDY
                     KLTRKQQEEEKRRYRDQKKMQDLLIKRRESIYGSKPSPRRSNSVRKTNGYNGDASVPP
                     TPRRNSAGATNNDIMTTPRSYSSHRQNGYFKEVRRLSTAPLNFVAIPKEDSVSTYTSV
                     CGSEPDSPLYN"
     CDS             join(26661. .26948,27454. .27610,27713. .27814,27911. .28038,
                     28342. .28470,28565. .28632,28747. .28969)
                     /note="similar to acid phosphatase type 5 emb|CAB63938.1;
```

```
                similar to EST gb|N96048.1"
                /codon_start=1
                /evidence=not_experimental
                /product="F10B6.10"
                /protein_id="AAF79221.1"
                /db_xref="GI:8778212"
                /translation="MTYIYRDTKITTKSTIPFLIFFLFCFSNLSMATLKHKPVNLVFY
                VYNLIIIFSSHSSTAELRRLLQPSKTDGTVSFLVIGDWGRRGSYNQSQVALQMGEIGE

  Query Match             7.7%;  Score 133.6;  DB 12;  Length 132699;
  Best Local Similarity   52.2%;  Pred. No. 2.3e-18;
  Matches  448;  Conservative    0;  Mismatches  389;  Indels   21;  Gaps    6;

Qy     560 cctttcataaggtttactcaattaaccgctaatcaagcgattttagaagcgattaacggt 619
           ||||   | | |||  |||| || || || ||||||||||||| | |||||  || | ||
Db 118136 CCTTATCTCAAGTTCGCTCACTTCACGGCGAATCAAGCGATTCTCGAAGCTTTTCAAGGG 118195

Qy     620 aatcatcaagcaatccacatcgttgatttcgacattaatcacggggttcaatggccaccg 679
           ||  |   ||   | ||  || ||||||||   || | ||| ||  ||||||||||  ||
Db 118196 AAGAAAAGAG---TTCATGTCATTGATTTCTCTATGAGTCAAGGTCTTCAATGGCCGGCG 118252

Qy     680 ttaatgcaagcactagctgatcgttaccctgctcccactctt---cgaatcaccggtact 736
           | ||||| || || ||   |||      || |||  || ||   ||  | ||||| | |
Db 118253 CTTATGCAGGCTCTTGCGCTTCGACCTGGTGGTCCTCCTGTTTTCCGGTTAACCGGAATT 118312

Qy     737 gg---aaatgaccttgataccttcgtagaacaggtgatcgtttagctaaatttgctcac 793
           ||   |   | |  ||||   |  |       |||   |     |||  | |||||
Db 118313 GGTCCACCGGCACCGGATAATTTCGATTATCTTCATGAAGTTGGGTGTAAGCTGGCTCAT 118372

Qy     794 tcattagggttgagatttcaattccatcctctttatatagccaataataaccacgatcac 853
           | |   | |  ||   |||| |  |   |   || | ||    |       ||  |
Db 118373 TTAGCTGAGGCGA---TTCACGTTGAGTTTGAGTACAGAGGATTTGTGGCTAACACTTTA 118429

Qy     854 gatgaagatccttctattatttcctccattgtactactccctgatgaaaccctagctatc 913
           | |||   |  | ||   ||        ||  || |        |||| |  | ||  |
Db 118430 GCTGATCTTGATGCTTCGATGCTTGAGCTTAGACCAAGTGAGATTGAATCTGTTGCGGTT 118489

Qy     914 aactgtgttttctacctccaccgccttttaaaagaccgcgaaaagttaaggatttttttg 973
           |||| |||||| | || |||   || ||   |   | |    | |    |   || |
Db 118490 AACTCTGTTTTCGAGCTTCACAAGCTCTTGGGACGACCTGGTGCGATCGATAAGGTTCTT 118549

Qy     974 catagggttaagtcaatgaaccctaaaattgttacaatcgcggagaaggaagcaaatcat 1033
             | ||| ||    || || ||  | ||| | ||  | |  ||| ||||| | || |||
Db 118550 GGTGTGGTGAATCAGATTAAACCGGAGATTTTCACTGTGGTTGAGCAGGAATCGAACCAT 118609

Qy    1034 aaccatcctctttttttacaaagattcatcgaggcgttggattattatacagctgtgttt 1093
           ||   |||  |||| ||| |  | || |  ||| ||||| |||||||  |  |  |||||
Db 118610 AATAGTCCGATTTTCTTAGATCGGTTTACTGAGTCGTTGCATTATTACTCGACGTTGTTT 118669

Qy    1094 gattcactggaagctacattgccaccgggtagtcgagagaggatgacagttgaacaagtg 1153
           || ||  |||||| |       |||| || ||| ||| | | | |   |     | ||
Db 118670 GACTCGTTGGAAGGT------GTACCGAGTGGTCAAGACAAGGTCA---TGTCGGAGGTT 118720

Qy    1154 tggtttgggagagagattgttgatatcgttgcgatggaaggagataaaaggaaagaaaga 1213
           |  || || | | ||||     |  | || ||    || |||  | |  |    ||  |
Db 118721 TACTTGGGTAAACAGATCTGCAACGTTGTGGCTTGTGATGGACCTGACCGAGTTGAGCGT 118780

Qy    1214 catgaaaggtttagatcatgggaagttatgttgaggagttgtggatttagtaatgttgct 1273
           ||||||| ||| ||    |||        |||  ||  |  ||| |||     ||    |
Db 118781 CATGAAACGTTGAGTCAGTGGAGGAACCGGTTCGGGTCTGCTGGGTTTGCGGCTGCACAT 118840

Qy    1274 ttaagccctttgcattatcacaagctaagcttctttgagacttcattatccttctgaa 1333.
           |  |  |   ||| ||    ||||| |   | ||||||  ||     |       |||
Db 118841 ATTGGTTCGAATGCGTTTAAGCAAGCGAGTATGCTTTTGGCTCTGTTCAACGGCGGTGAG 118900

Qy    1334 ggctatcaactcggagtttcgagtaattctttcttcttaggttggcaaaatcaacccctt 1393
           || ||||   | |  |           | | || | || |||||||| |  | ||| ||
Db 118901 GGTTATCGGGTGGAGGAGAGTGACGGCTGTCTCATGTTGGGTTGGCACACACGACCGCTC 118960

Qy    1394 ttctccatctcgtcttgg 1411
           |  ||| |||| |||||
Db 118961 ATAGCCACCTCGGCTTGG 118978
```

Search completed: June 30, 2001, 05:28:45
Job time: 4329 sec